A domain-independent model of suspense in narrative

Thesis

Link(s) to article on publisher's website:
http://dx.doi.org/doi:10.21954/ou.ro.0000ab68

A domain-independent model of suspense in narrative

Richard Doust

Faculty of Mathematics, Computing and Technology
The Open University

A thesis submitted for the degree of

Doctor of Philosophy

March 2015

Suspense is the nervous system of drama...

Alfred Hennequin

'The Art of Playwriting'

(Hennequin, 1890)

Abstract

Many computational models of narrative have focussed on the structure of the narrative world. Such models have been implemented in a wide variety of systems, often linked to characters' goals and plans, where the goal of creating *suspenseful* stories is baked into the structure of each system. There is no portable, independently motivated idea of what makes a suspenseful story.

Our approach is instead to take the phenomenon of suspense as the starting point. We extend an existing psychological model of narrative by Brewer and Lichtenstein (1982) which postulates suspense, curiosity and surprise as the fundamental elements of entertaining stories. We build a formal model of these phenomena using structures we call narrative threads.

Narrative threads are a formal description of a reader's expectations about what might happen next in a given story. Our model uses a measure for the imminence of the predicted conflict between narrative threads to create a suspense profile for a given story. We also identify two types of suspense: **conflict-based** and **revelatory** suspense.

We tested the validity of our model by asking participants to give step-by-step self-reported suspense levels on reading online story variants. The results show that the normalised average scores of participants (N = 46) are in agreement with the values predicted by our model to a high level of statistical significance.

Our model's interface with storyworld knowledge is compatible with recent developments in automatic harvesting of world knowledge in the form of event chains such as Chambers and Jurafsky (2008). This means that it

is in principle scalable. By disentangling suspense from specific narrative content and planning strategies, we arrive at a domain-independent model that can be reused within different narrative generation systems. We see our work as a signpost to encourage the further development of narrative models based on what we see as its fundamental ingredients.

Acknowledgements

The story of this research goes back a long way. With the help of a M.Sc in Cognitive Science at Essex University, I had obtained a post of *wissenschaftlicher Mitarbeiter* (research assistant) for Professor Helmut Schnelle in the Linguistics Department of the Ruhr University at Bochum in Germany. We were working on a translation protocol of an Earley parser into a connectionist network to prove their formal equivalence (see Schnelle and Doust, 1988, 1992). Perhaps not surprisingly, there are some quite striking similarities between that research and the model of suspense I am proposing today...

The research was successful and led to a number of publications (Wilkens and Schnelle, 1990, 1994). I am for ever grateful to Prof. Schnelle for my time in his department. We had numerous exchanges on linguistic, philosophical and cognitive issues, and he was a very personable support to me. At that time in my life, however, I needed to attempt to scale the wall to understanding the human brain from another, more artistic side and launched myself not without some regrets into a multilingual, musical and theatrical career based around Strasbourg in France.

Many years later, I wanted to find a way to connect the experience of my artistic work to scientific practice and I got back in touch with Prof.

Schnelle in Berlin. With his help, I was able to get accepted on this doctoral research program. So my first tremendous thanks go to Helmut Schnelle for his wide-ranging and profoundly *menschlich* inspiration.

Above all, my heartfelt thanks and gratitude go to my main supervisors, Dr. Richard Power and Dr. Paul Piwek, for all their untiring support, subtle guidance, productive silences and numerous thoughtful insights that kept me on my way in the maze of intuitions, formalisms, distractions and breakthroughs that this research has produced over the last six years. I have learnt so much and I am so grateful for all our suspenseful exchanges.

I am also very grateful to Dr. Allan Third, who joined the supervision in the last phase, for his support and for accepting to be the third supervisor of this thesis, and to Prof. Donia Scott who was part of the team at the beginnings of this project and whose vision was vital to finding the spark of suspense. I also want to thank Prof. Dr. Marian Petre for her encouragement and inspiring training sessions and to all who organised the graduate conferences: they were absolutely essential to the success of this research. Thanks too go to Prof. Noel Sharkey who remembered my work in Essex University and helped with his glowing references.

Friendly thanks also go to Brian Plüss, Eva Banik and all the other Open University post-graduate students who were always ready to help out and answer my questions. Many thanks too to all those who took part in the online experiments.

I especially want to thank my parents and my whole wide-ranging family who have followed the ups and downs of this process and whose support I have always felt. This journey has had its difficult moments and taken up a lot of my time and energy. I want to thank with all my love, Isabelle Marx, my partner and my friend throughout, for all she has done to make it

possible.

This research is part of my answer to a long-held dream of mine to build bridges between the artistic and scientific worlds. I hope to continue this bridge-building for a long time to come.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 From stories to suspense

1.1.1 The power of stories

In 1944 a group of psychology researchers designed an experiment using an animated film in which circles, squares and triangles of different sizes moved around inside a larger square (Heider and Simmel, 1944). Participants were asked to watch the film and then describe what they had seen. Many described the movements of these geometric forms in the form of a story made up of events such as: 'the big triangle chases the small one, and then the small circle comes to aid of the small triangle'.

The results of this experiment suggest two distinct statements:

- Human beings have a strong capacity for personifying even very abstract objects
- Human beings tend to organise collections of events into story-like structures.

Narrative psychologists have claimed that human beings perceive and interpret activities by structuring them into stories and also have emphasised the role that stories play in the way people make sense of their lives and organise their experience and knowledge (Bruner, 1991, Kerby, 1991). The central role that stories play in learning because they are so easily understood and remembered, together with their potential to improve communication and change management in organisations has also been emphasised by Lämsä and Sintonen (2006). Snowden (2000) claims that it is easier and more natural to use narratives than written knowledge to store information, and proposes the construction of 'narrative databases'.

The strong story claim is to say that *all* our thoughts and experience organise themselves into stories, that this is how human beings retain and recall events, and to a certain extent, even how they understand the events of their own lives. Even if this claim may seem exaggerated, what is clear is that in everyday life, stories are pervasive, and that human beings have cognitive abilities which are well-tuned to creating and understanding them.

1.1.2 Why are stories popular?

Hasson et al. (2008) and other researchers examined participants using a fMRI brain-scanning device while they watched film scenes. They claim that a certain Hitchcock episode triggered highly similar responses in more than 65 percent of the neocortex of the participants. Hasson's group also measured the participants' gaze during the viewing and the gaze maps were almost identical. These findings suggest a neural correlate for the claim that story-telling in the form of a film can exert a high degree of control over the attention of the spectators.

Further, Csikszentmihalyi and Csikszentmihalyi (1991) describes the complete absorption we can experience when we perform certain activities. Included in these activities cited of course is the reading or viewing of a story. It seems that novels and films are purposefully *designed* to produce this absorption effect (called 'flow'), and the pleasurable feeling that accompanies it. Suspense is one narrative phenomenon capable of maintaining the attention of viewers or readers over surprisingly long periods of time. Csíkszentmihályi suggests that through suspense's capacity to focus attention, it may strengthen the effects of the emotions that are experienced during the reading or viewing of a story.

At least one study (Abuhamdeh et al., 2015) shows that video games provoking higher uncertainty levels (and thus perhaps also higher suspense levels) were preferred by players even if it meant choosing games in which they had lower competence.

So, the absorption and extra attention that suspense triggers may be one of the factors in its success. But how do stories trigger these suspenseful moments? Just how do narrative structures such as a Hitchcock film generate the well-known feeling of suspense?

1.1.3 The suspense reaction: an evolutionary account

Narrative in general has been studied for its adaptive evolutionary function for our species: Tooby and Cosmides (2001) examines the possible adaptive nature of fictional narratives in human societies, Sugiyama (2001a,b) explores how foraging information can be packed into narrative form and Boyd (2005) discusses the different functions of art and narrative in evolutionary terms. In order to suggest a link from our research to this work, we propose a short evolutionary account of how suspense is evoked by the telling of a story. We call this phenomenon the *suspense reaction*. The following is not part of our main claim but aims to sketch a plausible general context for our approach.

The fight-or-flight response

Part of the suspense reaction may be linked to the well-known fight-or-flight response, first described by Walter Cannon (Cannon, 1932). Under this theory, animals react to threats by a reaction of the entire sympathetic nervous system which primes the animal to either fight or flee. Jansen et al. (1995) has detailed the role of central command neurons in triggering this cascading reaction which has physiological, emotional and cognitive components. Emotional arousal increases together with certain bodily responses including a faster heart rate, higher blood pressure, and vasoconstriction. De Wied (1995) calls this an 'anticipatory stress reaction'. As Schauer and Elbert (2010) says, 'evolution has equipped us with a defense armament to imminent threat'. Importantly, Schauer and Elbert extends Cannon's initial fight-or-flight reaction to a 'freeze-flight-fight-fright-flag-faint' reaction. These different reactions provide 'optimal adaption for particular stages of imminence'. 'Freezing' is called for when there is a large distance between the subject and the threat, 'flight' when the distance is reduced and 'fight' when the distance is eliminated. Physiological reactions thus increase by degrees according to the reduction in the perceived distance from the source of the threat. For the moment, we note that different degrees of imminence provoke differentiated bodily responses[1].

The stress of not knowing

Our starting assumption for this research is the idea that suspense is triggered by and related to the degree of stress and/or arousal we feel because we **do**

[1] We will refer again to this relationship between bodily reactions and perceived distance in our discussion of imminence in 3.3.3.

not know something we need to know about an ongoing process. Many people can experience a degree of bodily tension when experiencing situations such as say a football match, or narratives such as a film. In this tense, uncomfortable state, a great deal of our cognitive activity is focussed on detecting anything in the incoming information which could help us to regain clear predictions of future events. The ability to be able to predict what will happen next in a given situation may have been so important in evolutionary terms that we developed a reaction of heightened preparedness and attention, leading even to physical tension, in situations where we do *not* know what will happen next. If this is the case, we suggest that the suspense reaction could be both a kind of emotional and cognitive *preparation* for a potential fight-or-flight response and also potentially *part* of such a response.

Suspense as an essential ingredient in narrative

We further suggest that stories, amongst other things, are *parasitical* on the suspense reaction and purposefully provoke and maintain it, often many times during their telling. This view links to evolution-based accounts of cultural artefacts (see Tooby and Cosmides, 2001) which suggest that we are addicted to stories for reasons which have to do with our evolved survival instincts and that, in a similar way to humour and music, stories simulate, exercise, train or exploit cognitive and emotional tasks that were essential for survival in the world in which mankind evolved.

1.1.4 Why does clarifying the concept of suspense matter?

Aside from a very frequent use of the word to describe what we feel while watching for example, a Hitchcock film, the word 'suspense' can also be used to describe the tension in a very close sports match, or even the rather nebulous waiting for an exam result. The fundamental metaphor at work suggests that something, perhaps a decision or an action, is 'hanging in the air' and has not yet 'fallen to the ground' and become part of common knowledge.

Dictionary definitions of the word 'suspense' suggest that the word is more like a concept cluster than one single well-defined concept. The English dictionary (Collins, 2003) gives three definitions:

1. apprehension about what is going to happen.
2. an uncertain cognitive state; 'the matter remained in suspense for several years'
3. excited anticipation of an approaching climax; 'the play kept the audience in suspense', anticipation, expectancy - an expectation.

Generally, it seems that an overarching and precise definition of suspense to connect the above meanings is lacking. One goal of this research is to create a model of suspense that can throw some light on all three of the above definitions. We will exclude uses of the word which appear to be different to these meanings, and we will be looking for a more fundamental definition which reconnects the variety of definitions present in scientific literature.

A central motivation for this research is that if we can explain suspense then we can create suspenseful stories. A standardised procedure for measuring, comparing and controlling suspense could be used in the following domains of application:

Natural language and story generation: Starting from basic story-lines, the addition of a suspense module to a wide range of interactive and non-interactive narrative systems could provide new ways to create engaging text-based and film-based stories.

Performing arts: New analytical teaching tools for creative writing, film-making, play-writing, stage performance and musical composition could be developed. We could explain part of the entertainment value of a story in terms of a formal analysis of its suspensefulness. Theories about the structure of for example the fairy-tale (see Propp, 1968), could be seen as particular instantiations of a set of fundamental narrative mechanisms.

1.2 Constraining the research

1.2.1 What kinds of story are we dealing with?

A narrative consists of a sequence of events communicated from a particular perspective. Some narratives such as reports, describe *processes* and have an explanatory function. Other narratives have the function of *entertaining*, by provoking different types of emotional and cognitive engagement from the reader. The word 'narrative' comes from the Latin for telling a story (*narratus*) and the word has been used for both the actual *physical realisation* of the telling of a story, such as a text or film (Genette, 1972) and as a synonym for *story* itself (Barthes, 1966). Following Brewer and Lichtenstein (1982), we take a story to be a narrative which has entertainment as its main *raison d'être* and also some kind of closure and internal coherence.

Apart from suspense, stories evoke other emotional and cognitive reactions such as surprise and curiosity, and different story genres can emphasise one phenomenon more than others. Suspense might be said to dominate in the Western or thriller genres and curiosity in the detective story. For this research, we will only be concerning ourselves with very short stories which contain at least one suspenseful situation.

Stories often use techniques such as reported events in speech, past and future perfect tense, changing points of view, flashbacks and flashforwards. Such techniques typically change the order in which the story releases or reveals information to the reader. As narrative techniques, they can, of course, have a strong effect on the suspense of a story. For this research, we will exclude such techniques and assume that we are dealing with **chronological stories**, that is, stories which describe sequences of events in their chronological (and usually also causal) ordering.

Our goal is to create as much clarity in our formulation as possible; it is therefore important to restrict our domain of application. We suggest that once we have a robust model of suspense that can deal with chronological stories, the effect of the above narrative techniques will be much easier to understand.

1.2.2 Brewer and Lichtenstein's model of suspense

The background to our research comes from the structural affect theory developed in Brewer and Lichtenstein (1982). This theory shows how a sequence of events can be told in three distinct ways to create suspense, surprise and curiosity. They give the following short story as an example of a story that produces suspense:

> The sniper was waiting outside the house. Charles got up from the chair. He walked slowly toward the window. There was the sound of a shot and the window broke. Charles fell dead.

According to Brewer and Lichtenstein, this sequence has an *Initiating Event* which introduces the sniper and an *Outcome Event* which narrates Charles' demise. To create surprise from the same sequence of events, we need only leave out the first *Initiating Event*, as Brewer and Lichtenstein show in the following example:

> Charles got up from the chair. He walked slowly toward the window. The window broke and Charles fell dead. The sound of a shot echoed in the distance.

Lastly, to produce curiosity, an event must also be left out, but in such a way that the reader *knows* that something is missing. To illustrate this, Brewer and Lichtenstein give the following story:

> Charles fell dead. The police came and found the broken glass, etc.

This concept of Initiating and Outcome Events will play an important role in the development of our model of suspense. We will build on the idea that an Initiating Event triggers the prediction of an Outcome Event. In addition, we use the concept of a *conflict* or an *incompatibility* between certain predicted *Outcome Events* to characterise the suspense in a story.

1.2.3 Imminence

The term 'suspense' can be used for someone waiting for an exam result, for example. In some sense, the suspense in this situation remains exactly the same whether we are three weeks or 10 minutes away from the result coming out. But, of course, we know that the emotions felt in the second situation are much stronger. It is as if the danger were *closer* and therefore *greater*. To capture this effect, we will add an additional time-based feature called 'imminence' to our categorisation of suspenseful situations.

In this way, we separate conceptually the actual *conflict* between predicted Outcome Events in a story from the *imminence of the resolution* of such a conflict. Our claim is that quite often the same basic conflict between predicted events can be present at the beginning of an episode in a story as at the end; all that changes during the telling is the likelihood that the conflict will be decided one way or the other *within the next few story steps*. We therefore distinguish the *type* of interaction between events from the predicted *time* of their interaction.

Of course, for stories, we are dealing with imminence in terms of the actual *telling time of the story*, and not necessarily an actual length of time in the world in which the story occurs. An event which might be a year away in the story could conceivably be perceived as highly imminent during the telling of the story. This is indeed one of the strengths of stories as compared to models of real-life situations; they have a capacity to concentrate on what is interesting and speed up and slow down time as needed.

1.2.4 The goal of our work

The main goal of this research is to find a general way to formally model suspense in stories. Of course, readers may have their own idiosyncratic suspense reactions while reading a given story. We will start however with the assumption that each story has an identifiable generic *suspense profile* which can be represented by a curve of suspense values for each step in the story as the story is told. Such a suspense profile might correspond to the averages of the perceived suspense values for a large population of readers of the story. Our model should therefore be able to explain and derive the fluctuation of suspense values over different parts of a story, or in other words, determine its *suspense profile*.

Our model of suspense will be based on the structure of the flow between the different inferences that a story triggers. However, our approach will be to propose a way of capturing story-relevant inferences which is not formally dependent on specific models of agent behaviour such as goal and planning models.

1.2.5 Our research question

Our research question can be formulated in the following way:

- What are the key components of a formal model of suspense that allows us to correctly measure and control suspense in narrative, whilst using a generic, domain-independent model of the story content?

Our model should have some psychologically plausibility and yet be computationally tractable. Ideally, we want to be able to track the suspense felt by a reader[2] step by step as a story is told.

1.3 Contributions of this research

In order to present this research succinctly, we will often use the term 'storyworld'. We define this as the sum of all the information which is necessary to fully understand a series of related stories. The storyworld is intuitively equivalent to the *setting* in which a given story occurs, together with all the *causal and intentional rules* which govern the different possible events that can occur in it.

In answer to our research question, we will make the following contributions:

- A **formal domain-independent model of suspense** based on calculating the predicted conflict between narrative threads which extends

[2]We will mostly use the word 'reader', but we will assume that the descriptive level we are dealing with can be applied with equal success to spectators of drama as well as to readers and listeners of stories.

the work of Brewer and Lichtenstein (1982). We extend the concept of Initiating and Outcome Events to a list-like structure which we call a **narrative thread** and which is independent of particular ways to model story information. This allows us to model the predicted conflicts between different outcomes that are triggered by a story in terms of mutually incompatible events in narrative threads. We then model suspense using intermediate variables of Imminence, Importance, Foregrounding and Confidence, and we present a full mathematical description of our narrative thread model.

- A **method for creating a computational model of a storyworld** which depends on causal and intentional storyworld information. We start by considering certain features of the storyworld in which a given story is set. We then use general criteria to build up a set of narrative threads and a set of mutually incompatible event-pairs to encode the storyworld.

- A **computational implementation of our suspense model** which uses storyworld information to derive the suspense profile of a given story. The implementation of our model uses a number of fixed internal parameters which regulate the relative effects of the intermediate variables. It also requires a degree of calibration to determine the relative importance of the different narrative threads.

- An **experimental method for measuring people's suspense profiles** as they read a story which will enable us to test our model's predictions. This uses a free-scale magnitude estimation method based on step-by-step self-reported suspense ratings. We use the averaged suspense ratings from a group of participants in order to calibrate our suspense implementation and also to test the predictions of our model on a different story-variant from the same storyworld.

1.4 Plan of the thesis

This thesis contains the following chapters:

- Chapter 2: A review of the concepts of suspense which occur in the psychological, literary and computational literatures
- Chapter 3: An non-formal argument building up our model of suspense
- Chapter 4: A mathematical formulation of our suspense model
- Chapter 5: A description of a computational implementation of our model applied to a simple short story
- Chapter 6: An online experiment designed to test the implementation with a story-variant
- Chapter 7: A summary of our main conclusions together with suggestions for future work

Chapter 2

"What is suspense anyway?": a review of literature on suspense

A perusal of literature claiming to teach story structure and plot reveals surprisingly little mention of suspense. In 'The Anatomy of Story' (Truby, 2007), a major work on how to become a master story-teller, the word 'suspense' does not even occur once in 445 pages. Dibell's book 'Plot' (Dibell, 1988) however, does mention suspense, and suggests three ways to produce it:

1. Switching plots, that is, using subplots to slow the main action. (ibid.,p.63)
2. 'Waiting to find out builds suspense, drama' (ibid.,p.89).
3. Using the 'Rule of three', that is, three repetitions of an event to heighten the expectations about the outcome and

how the outcome might be different this time. (ibid.,p.89)

It would seem that the information which is designed to help the would-be writer concentrates mainly on what *needs to be added* to already present basic human intuitions about characters and plot to *improve* the story telling experience. Accordingly, such approaches leave out many details and take much for granted. It seems that the authors of such books have an intuitive grasp of suspense and presume that so too do their readers.

We will be able to reexamine the methods proposed by Dibell above in the light of our suspense model in our conclusions. For the moment, however, we take note of a relative dearth of information about suspense in the story-writing paradigm. This suggests that techniques for producing suspense remain mostly at an intuitive level and perhaps also that suspense is for the moment difficult to talk about in a precise way. As an example of this, in popular parlance, suspense is sometimes considered equivalent to a cognitive state of uncertainty whilst also describing the emotional reaction of anxiety that such a cognitive state produces. There is clearly a need to disentangle effects from causes. What is missing is a clear recipe for creating and maintaining suspense in a story which would explain for example the three methods above.

2.1 Suspense as an object of scientific research

Stories are so ubiquitous in human activity that scientific approaches to understanding them appear in a range of academic disciplines. Consequently there are also a range of different fields in which the term ‘suspense’ appears as a subject of scientific scrutiny. Our goal here is to provide the necessary context to our model of the phenomenon of suspense. In this light, we will be focussing essentially on the *relationship* of various models of narrative to suspense. We will be distinguishing scientific work on narrative which occurred before the advent of the computational paradigm in scientific work or which were little influenced by it, from other work which draws more or less on the new possibilities that this paradigm offers.

We will structure our review in the following way:

- Pre-computational work on narrative, including literary and aesthetic theories of narrative
- Psychological approaches to narrative comprehension
- Psychological approaches to suspense
- Computational models of narrative and their relation to suspense

2.2 Pre-computational and literary theories of narrative with regard to suspense

2.2.1 Introduction

Pre-computational views on narrative have much in common. One common characteristic is that they do not make explicit a theory of narrative comprehension, concentrating instead on the structure of the narrative itself. In different ways and for different reasons, many concern themselves with the idea of a *plot*. Plot is strongly linked to suspense and the manipulation of a plot can be construed as one of the ways that suspenseful narratives are created. Showing how plot as a concept has developed will bring up other useful concepts which will help the development of our suspense model. We start our brief overview with Aristotle and end with some 20th century literature theorists.

2.2.2 Plot and conflict

Around 330 B.C., the Aristotelian concept of a plot or a *mythos* as it appears in his Poetics (Aristotle, 1974, X-XI), referred to the structure of the incidents in a story and was one of the essential components of tragedy. A good complex plot had to have a 'Reversal of the Situation', that is, a change of fortune from good to bad or from bad to good and also a moment of 'Recognition', that is, a change from ignorance to knowledge. Also, the plot of a story had to have a beginning, a middle, and an end. Very roughly, the beginning is where the characters are presented and the conflict is initiated, the middle is where the conflict develops, and the last part is where the conflict is resolved (ibid., VII).

In 1863, over 2000 years after Aristotle's Poetics, Freytag refined Aristotle's theory of tragedy and plot and created the 'Freytag pyramid' in his 1876 work *Die Technik des Dramas* (Freytag, 1863). In this model, a dramatic work can be split into five functional parts:

1. Exposition
2. Rising action (through conflict)
3. Climax
4. Falling action
5. Resolution.

After the initial *Exposition* phase in which the main characters of the story appear, comes an *Inciting Incident* which starts off the conflict that defines the *Rising Action* phase. This phase leads up to the *Climax* where decisive, story-defining actions occur. The *Falling action* phase consists of the resolution of the story protagonists' main problems and leads to the final *Resolution*.

The *Rising action* phase is where the main conflict in the story occurs. The concept of conflict contains an idea of uncertainty of outcome and is a key part of our suspense model as we shall see.

2.2.3 Plot and story

The Russian Formalists

In the early 20th century, some fifty years after Freytag's Pyramid, the influential school of Russian formalists extended the concept of *plot*, creating new concepts for the analysis of narrative.

One important distinction they introduced was between Фабула (*Fabula*) and Сюжет (*Syuzhet*)[1]. The *Fabula* is the actual chronological sequence of causally-related events in a given storyworld, whereas the *Syuzhet* is the *way* in which this sequence of events was actually revealed and manipulated in the telling of the story. Thus, to create the *Syuzhet*, the events in the *Fabula* could be omitted, delayed or told in a non-chronological order by means of flashbacks and flashforwards.

This distinction has proven useful in many models of narrative and is also part of our formalisation of suspense.

Propp's recipe

One of the Russian formalists, Vladimir Propp developed a kind of formal recipe for story creation. In his 'Morphology of the Folktale', Propp (1968)

[1] *Fabula* and *Syuzhet* were developed amongst others by Vladimir Propp (1928), tr. Propp (1968) and Shklovsky (1917), tr. Shklovsky (1965). Both Russian words come from Latin: *Fabula* means 'fable' and *Syuzhet* (pronounced *'syougette'*), comes via the French 'sujet' from 'subjectus'.

suggests that a limited number of narrative situations can be used to characterise almost any folk-tale and also many other types of story. These situations include the following: 'Receipt of a magical agent, Guidance, Struggle, Victory, Return, Pursuit, Rescue'.

We can see these elements as typical ways to instantiate the different parts of Freytag's Pyramid; the 'Receipt of a magical agent' or 'Guidance' could be *Inciting Incidents* which start the conflict phase, 'Struggle' and 'Pursuit' can be seen as typical *Rising Actions*, that is, conflict-based suspense producing phases and 'Victory' and perhaps 'Return' examples of *Falling Actions* or *Resolutions*. These situations have the advantage of both fitting into the Pyramid and evoking strong empathetical reactions in the listener. The notions of empathy and conflict will both be essential to our account of suspense.

Transforming a Fabula into a Syuzhet

Grd Genette developed a taxonomy of classical narratology based on the work of the Russian formalists. In his 'Narrative discourse', Genette (1972) describes three ways in which a *Fabula* can be transformed into a *Syuzhet*:

Tense: the way events are placed in time and delayed, repeated or ordered.

Mood: the emotional relationship of the narrator to the events as they are presented.

Voice: the choice of the narrator to relate the events.

This analysis emphasises the need to make choices, rather like applying different filters to the underlying causally-related events of the *Fabula*, in order to produce a particular version of the story.

Character and spectator knowledge about events in the Fabula

An oft quoted necessary condition for suspenseful drama is a *lack of important information*. As Hitchcock (1956) says, 'The audience knows that a given piece of information is missing, but does not know what it is.' This feature would, however, perhaps be better described as triggering curiosity, rather than suspense. Of course, as most suspenseful narratives also seem to use in one way or another the notion of missing information, it does indeed seem that the notions of curiosity and suspense are strongly linked. In this regard, White (1939) had already claimed that suspense is 'prolonged curiosity'. Hitchcock's view of suspense as lack of information highlights once more the difficulty of distinguishing these concepts and in making clear how they work together in narrative.

Bal (1997, p.114) and Branigan (1996, p. 75) from the field of asthetics and literature theory, formalised a typology of possible relationships between the reader and characters in narratives. To distinguish the different narrative structures, they imagine asking questions of both reader and characters and determine which of the latter would know the answers. The four different cases they came up with can be summarised as follows:

1. spectator does not know and character does not know
   ⇒ riddles, detective stories: suspense is present
2. spectator knows and character does not know
   ⇒ thriller stories: suspense is present
3. spectator does not know and character knows
   ⇒ 'secret' stories: suspense is present
4. spectator knows and character knows
   ⇒ no suspense is present

We can see that under this classification, three structurally different ways to produce suspense can be distinguished. However, from the point of view of the *spectator*, there are in fact only two: i) the spectator knows and the character does not know, and ii) the spectator does not know. These two suspense types will be dealt with by our model.

2.2.4 The nervous system of drama

We end this overview of pre-computational theories of suspense with a brief presentation of a treatise by Alfred Hennequin. Hennequin (1890) introduced seven means for maintaining interest in a play in his *The Art of Playwriting*. One of these was suspense which he called the 'nervous system of drama'. His account of suspense is mostly based on the listener's doubt about what will happen next. Some of his insights are:

- Suspense can still exist even when the author appears to show us exactly what is going to happen.
- If one element of suspense is removed then it should be replaced by another, and this can be done by the introduction of an additional 'obstacle'. This can be done in four ways:
    1. By interposing some new and unexpected obstacle.
    2. By emphasising some obstacle already known to exist.
    3. By bringing to light an obstacle which is at once seen to have existed all the time.
    4. By causing a new obstacle to result from the very removal of others.

Hennequin's obstacle technique for producing suspense can be seen as a way of introducing an element of *conflict* into a situation but also of *delaying* *the resolution* of a conflictual situation.

2.3 Psychological theories of narrative comprehension

2.3.1 Introduction

Psychological approaches to narrative focus a great deal of attention on real-life processes, often including theoretical descriptions of the narrative comprehension process. As the object of our concern, suspense, is a real-time reaction to the unfolding comprehension of a story, we will look at both:

- Psychological theories of narrative comprehension, and
- Psychological theories of suspense.

Our approach to narrative modelling is designed to be medium-independent; we hope to provide a model of suspense which is valid for narratives using text and still or moving images. For this reason, we will not concern ourselves with basic text comprehension or image decomposition, but rather analyse the stages of comprehension that follow these preliminary processes.

2.3.2 Narrative comprehension processes

In addition to suggesting some important features of suspenseful situations which our model should take into account, an analysis of the assumptions used in psychological approaches to narrative comprehension can provide insights into the types of events and inferences that we should include in our model.

Kintsch (1988), Nathan et al. (1992), Van Dijk et al. (1983) have suggested that three levels of code are constructed in the textual narrative comprehension process. They use the following terms:

- the *surface code* which corresponds to the exact wording and syntax of the text,
- the *textbase*; this corresponds to the 'internal' and 'external' inferences needed to make the story coherent, and
- the *situation model*; this corresponds to some kind of 'mental description' of the events in the storyworld

Magliano et al. (2013) makes similar distinctions between front-end and back-end processing of narrative media and use the following similar terms for the different stages of narrative comprehension:

- Event segmentation
- Inferencing
- Structure building

We now examine research in these three areas with a view to its relevance to the question of suspense and then briefly discuss the role of emotion and narrative immersion.

Event segmentation

The first 'front-end' process in narrative comprehension that we will consider is event segmentation. Many studies show that people observing human activities will split them up into segments to enable better understanding and recall.

Firstly, research by Zacks et al. (2007) has identified a network of brain regions that are activated at event part boundaries, whether the participant is consciously attending to these boundaries or not. This study shows that people watching sequences of events spontaneously encode what they see in terms of their time-based parts and subparts.

Speer et al. (2007) has shown that a similar process also occurs in reading. Tests using fMRI showed that participants' neural activity increased at points in a narrative which corresponded to event boundaries. The texts used in this experiment were based on narrations of mundane everyday events. According to Speer, her work shows 'not only that readers are able to identify the structure of narrated activities, but also that this process of segmenting continuous text into discrete events occurs during normal reading' (ibid., p.3). The brain regions that responded were also the same as those activated when people viewed films of everyday events, and Speer suggests that this 'may reflect the existence of a general network for understanding event structure' (ibid., p.4). Recent work by Magliano et al. (2012) has also found very strong convergence on event boundary judgments across film and textual media.

Zwaan et al. (1995) and Zwaan and Radvansky (1998) propose an event-indexing model which lists features such as space, time, causality, and the goal episode that readers might use to update their comprehension of a text. The model can be understood as a protocol for a list of questions to ask of each new clause in a text-based narrative to determine whether a new event has occurred, as follows:

New time: is there a new temporal reference?

New space: has there been a spatial change?

New interaction: has a character changed their interaction with an object?

New subject: is there a new subject?

New cause: is any new activity not *directly caused* by previous activity?

New goal: has a character begun a new goal-directed activity?

In most psychological approaches to narrative, event segmentation appears to be an essential part of narrative comprehension. A complete and rigorous treatment of this phenomena lies however outside the scope of this research. We will content ourselves here with the preceding reformulation of Zwaan's protocol, which will serve to guide the segmentation of the stories used in the development of our model.

Attention span Studies on attention span (see for example Middendorf and Kalish, 1996) agree that people have two types of attention span:

- A *short* attention span that allows a response to events that last seconds.
- A *long* attention span that is a kind of sustained effort allowing the production of consistent results on some task over a time-scale of up to 20 minutes.

In a given narrative, sometimes many events can occur in a short period of time, thereby making strong demands on the *short* attention span. Conversely, sometimes very few events might occur over a much longer period of time, thereby making stronger demands on the *long* attention span. The question of event segmentation is therefore also linked to attention span.

Film directors create narratives in which the reader's temporal experience can be very precisely controlled and they can play with these two types of attention span to produce different effects. For example, in the culminating scene of the film 'The man who knew too much' (Hitchcock, 1956), Hitchcock maintains suspense over a long 12 minute sequence where very few events occur: the hero is trying to find the killer somewhere in the concert hall. Then at the climax, three dramatic events occur in the space of one second: there is a lull in the music, a woman screams and a shot is fired.

In this research, however, we will only be looking at very short stories and we will not go further into this question.

Explanation- and expectation-driven inference

The next intermediate phase of narrative comprehension is inferencing. Inferences are often split into two types (see Bower and Morrow (1990), McNamara and Magliano (2009), Singer and Ferreira (1983), Trabasso and Suh (1993)):

Internal inferences: These are directly available from the narrative and are based on relationships between explicitly mentioned narrative elements. For text-based narratives, for example, a new clause could be directly related to a previous clause in the text.

External inferences: Such inferences rely on general or specific knowledge structures available to the reader.

We now briefly examine the *Constructionist* and the *Prediction - substantiation* models of the processes underlying narrative comprehension:

The Constructionist model This model of narrative comprehension assumes that readers have specific goals in reading a text and constantly make an effort to 'search for meaning' as they read (see Graesser et al., 1994, for an in-depth presentation). In an analogy with computational on-line and off-line processes, the theory predicts that certain types of inference are made directly *as a text is being read and as a part of the reading process*, that is, they occur 'on-line', whereas others occur separately from reading, that is, they occur 'off-line'.

Off-line processes are also generally related to a degree of awareness of the units involved (words in linguistics for example), whereas on-line processes are considered to be unconscious. See for example Veldhuis and Kurvers (2012) for a discussion of the on-line/off-line continuum in relation to language segmentation.

We can readily imagine that due to additional conscious cognitive activity, certain off-line processes could in themselves evoke or increase the suspense felt by the reader of a story. An example might be a reader doing a kind of 'extra' independent worrying about a story situation. However, our goal in this research is to determine just how well suspense can be modelled using purely on-line processes.

Graesser et al. (ibid., p.4) provides a list of 13 **inference classes** that can occur during narrative comprehension. Guided by the constructionist principles that readers' main goals are to attempt to answer why-questions about events in the narrative whilst maintaining the global and local coherence of their understanding of the text, the constructionist theory notably predicts that of these inference classes, the following will *not* be generated on-line:

- Instantiation of a noun category
- Instrument
- Subordinate goal or action
- State

- Most cases of causal consequence inferences

These categories have in common that they are all concerned with 'filling in the details' of events. The instantiation of a noun category would require that the reader infer the existence of a *specific* meal when the word 'meal' is mentioned in a text, or similarly, a *man* or a *woman* when the word 'person' is mentioned. Subordinate goals or actions would be inferred in a similar way: the phrase 'she locked the door' would create the inference that a key was used. According to constructionist theory, under normal conditions of narrative comprehension this filling in of the details is *not performed*. The theoretical justification for this prediction is that such inferences are not needed to construct a *coherent* explanation of the narrative content.

In the context of suspense modelling, it is relevant to note that, according to the constructionist theory, only a very *specific* subset of causal consequence inferences can be constructed on-line:

- superordinate goals of existing plans,
- emotional reactions of characters
- causal consequences activated by several information sources
- causal consequences highly constrained by context with few alternatives.

The Prediction-Substantiation model This model (also described in Graesser et al. (1994)) is a related model of narrative comprehension which claims that narrative comprehension is not only *explanation-driven*, but also *expectation-driven*.

The main distinction between the two models for our purposes concerns inferences of the **causal consequence** type. Crucially, the prediction-substantiation model *includes* causal consequence inferences in its model of the online processes involved in reading a story. It claims that readers *not only* generate predictions about future occurrences in the plot, but also that these predictions *guide the narrative comprehension process.*

We can now situate the inferential processes we claim are needed for suspense in relation to these models of narrative comprehension.

Causal consequence inferences According to Graesser et al. (1994), a causal consequence inference can occur within the framework of the Constructionist theory under the following conditions:

1. The inference is supported by *several information sources.* This could happen, for example, when information from Short Term Memory and Long Term Memory both lead to the same inferences.

2. The inference is *highly constrained by context* and there are few if any alternative consequences that could occur.

The first condition suggests that inferential support by many sources alone can bring the possibility of the inference to the awareness of the reader: a kind of 'inference by association'. (Graesser et al.'s model uses the notion of an 'activation threshold' which must be reached before the inference can take place.) We can summarise this case by saying:

- If several information sources A, B, C... all imply Z, then Z will probably occur in the narrative

An example satisfying the second condition would be the inference that an object will drop, if it rolls off a table. This kind of inference could be described as the most basic inference possible:

- A 'always' implies B

Here the inverted commas suggest remind us that this is not an *absolute* rule, even though the chances of a different outcome are very low. These two conditions can be seen to emphasise the ease with which knowledge structures can be retrieved from long-term memory: we have at least one of the following:

- a strongly **supporting context**, or
- a strongly **directive context**.

One way in which these two conditions can be satisfied is by the activation of knowledge structures such as scripts, (see for example Schank and Abelson, 1975). Importantly, such predictions do not specify all the details of the events which could occur.

We will discuss the relevance of this result to our suspense model in the next chapter.

Structure building

The final 'back-end' phase of narrative comprehension occurs when mental structures corresponding to the perceived events in the storyworld are constructed by the reader. The creation of internal representations of narrative events which are in some way similar to real events, is the end result of the whole narrative comprehension process.

This phase has been modelled in different ways. Trabasso et al. (1989) uses a causal network based on settings, goals, attempts and outcomes. In their model, goals create expectations and outcomes either confirm or annul them. Zwaan et al. (1995) proposes a *situation model* based on events and intentional actions. The model uses the following indices:

- Temporality,

- Spatiality,
- Protagonist,
- Causality and
- Intention

and the reader updates their situation model whenever there is new information about one of these indices.

Brewer and Lichtenstein (1982, p. 3) has underlined the importance of memory constraints in the production of structured representations of narrative events, showing that for stories interpreted in terms of a hierarchically organised plan schema, story actions higher up in the hierarchy were more easily recalled. This would suggest that readers create a hierarchical organisational structure to help memorise and recall a story.

Finally, Baggett (1979) carried out a study on text-based and visually based narratives which provides support for the view that back-end processes are surprisingly similar regardless of the modality of the experience.

The role of immersion, emotion and empathy

We have looked at the cognitive part of narrative comprehension which depends on the way the narrative is structured and processed. In 'Why anyone would read a story anyway', Kintsch (1980) categorised narrative interest into two kinds:

- cognitive interest, arising from a well-organised discourse structure,
- emotional interest, arising from the emotional context of the story.

Recent evidence suggests that language understanding is 'grounded', that is, it depends on the brain systems that we use for moving around and perceiving the world (see for example Zwaan et al., 2004). Conversely, language can provoke the internal simulation of such movements and perceptions leading readers in some way to *situate themselves inside the story-world.* There are various theories about how this immersion might take place. Here is a sample:

The pretend theory (Walton, 1978): the readers *pretend* that the events in the story are real and feel 'quasi-emotions'.

The illusion theory (see Tan and Fasting, 1996, p. 236): the readers consider the story events to be 'almost' real, as if they themselves were inside the story-world.

The thought theory (Carroll et al., 1990): the readers imagine an emotional situation and this is enough to provoke an emotion.

In practice, for the purposes of understanding the phenomenon of suspense, the differences between these theories are not highly significant. We can consider that the pretending that goes on in the pretend theory might be the same phenomenon creating the *partial* illusion that occurs in the illusion theory. Similarly, the thought theory encompasses both of these, and is just a way of stating that states of mind can provoke emotions. We surmise that the differences between these theories correspond in fact to the differing degrees to which readers are *aware* of their own storyworld immersion and not to different types of immersion *per se.*

2.3.3 Psychological theories of suspense

We now examine some of the different ways that psychological approaches deal with the concept of suspense, before looking more closely at the approach taken by Brewer and Lichtenstein (1982).

A variety of suspense concepts

Scientific definitions use a variety of concepts:

- 'hope and fear':
  - Tan and Diteweg (1996, p. 151): 'The experience of suspense involves an emotional response, a state of fearful apprehension. Fearful apprehension may be seen as a prospect-based emotion, a class of emotions including hope, fear, and others...'
  - The cognitive appraisal paradigm (Ortony and Clore, 1989, p. 131): 'We view suspense as involving a Hope emotion and a Fear emotion coupled with the cognitive state of uncertainty'.
  - Sternberg et al. (1978, p. 65): '... suspense derives from a lack of desired information concerning the outcome of a conflict that is to take place in the narrative future, a lack that involves a clash of hope and fear...'
- 'expected negative outcomes':
  - Vorderer et al. (2001, p. 344): 'In a typical drama situation, when the character's failure becomes likely, they may even feel empathetic stress, a rather negative emotional experience better known as suspense.'

  - De Wied et al. (1992, p. 325): 'Film suspense can be described as an anticipatory emotion, initiated by an event which sets up anticipations about a forthcoming (harmful) outcome event for one of the main characters.'
  - Carroll (1984, p. 72): '... suspense in film is a) an affective concomitant of an answering scene or event which b) has two logically opposed outcomes such that c) one is morally correct but unlikely and the other is evil and likely.'

- 'number of solutions':

  - Gerrig and Bernardo (1994) suggest that reading fiction involves constantly looking for solutions to the plot-based dilemmas faced by the characters in a story world. One of the suggestions which come out of this work is that suspense is greater the lower the number of solutions to the hero's current problem that can be found by the reader.

Other concepts also often get a mention:

- 'structure':

  - Alwitt (2002, p. 35): 'Suspense is a cognitive and emotional reaction of a viewer, listener, or reader that is evoked by structural characteristics of an unfolding dramatic narrative.'

- 'uncertainty'

  - Carroll et al. (1996, p. 84): 'Suspense, in general, is an emotional state. It is the emotional response that one has to situations in which an outcome that concerns one is uncertain... If I believe that an outcome that I care about is uncertain, then suspense is in order.'

- 'what is at stake'
  - Caplin and Leahy (2001, p. 73): '... we define suspense as the pleasure experienced immediately prior to the anticipated resolution of uncertainty, and posit that it is positively related (up to a point) to the amount that is at stake on the outcome of an event.'
- 'curiosity'
  - White (1939, p. 40): 'Suspense is a continuous state of ungratified curiosity, and so keeping up the suspense is a matter of prolonging such a state... Suspense, being sustained curiosity, prolongs the change of experience that curiosity provides from the un-inquisitive state that preceded curiosity.'

This rather broad collection of concepts of suspense reveals perhaps the somewhat confused state of current knowledge about suspense. We hope through our research to contribute a theoretical clarification of the concept. To show the path we intend to follow, we describe the relationship of our research to the preceding definitions as follows:

- 'hope and fear': these notions will be grouped in a single emotion scale which can be positive or negative and vary in degree.
- 'expected negative outcomes': outcomes will be classed as positive or negative according to the emotions they provoke and also given a certain

importance level. Expectation or prediction will be a cornerstone of our model.

- 'solution': this term occurring in Gerrig and Bernardo (1994) is linked to a way of modelling the reader's thought processes that we will not be using explicitly. We will instead be using a simpler approach which models the reader's *predictions*.
- 'structure': we propose some *structural characteristics of information flow* that can explain important aspects of suspense.
- 'uncertainty': in our prediction-based model, we start by making the simplifying assumption that all predictions are *equally* likely.
- 'what is at stake': again, this concept will be covered by the (emotional) importance that a story outcome has for the reader.
- 'curiosity': we will also explore ways in which our model can also be used to model a type of suspense based on curiosity.

Structural affect theory: Brewer and Lichtenstein's approach

Brewer and Lichtenstein (1982) proposes a theory of narrative understanding from a psychological perspective. As we showed with the 'sniper' story in 1.2.2, they suggest that there are three major discourse structures which account for the enjoyment of a large number of stories: surprise, curiosity and suspense. This approach requires the existence of an Initiating Event (*IE*) and an Outcome Event (*OE*) in a given narrative.

For surprise, according to Brewer and Lichtenstein, some critical information from the event structure, that is the *IE*, is left out and importantly, the reader does not know that this information is missing. The leaving-out of information is *inconspicuous*. When the reader is presented with *OE*, they feel surprise and extrapolate the *IE* in hindsight.

In the case of curiosity, some information, that is, the *IE*, is omitted, but the reader is given enough information to know that this information is missing. The leaving-out of information is *conspicuous*. The reader thus becomes curious about the missing information and gradually fills in their knowledge about the *IE*.

Finally, for suspense, an *IE* is presented which triggers the prediction of an *OE* which could lead to significant consequences for one of the characters in the narrative. The reader feels concern about the effect of the outcome on this character, and if this state is maintained over time, the feeling of suspense arises. As Brewer and Lichtenstein say:

> 'often additional discourse material is placed between the initiating event and the outcome event, to encourage the build up of suspense' (ibid., p.17).

In suspense, therefore, *IE* and *OE* are ordered chronologically and other events are placed between them.

Sternberg's formulation of narrative dynamics Brewer and Lichtenstein's approach to suspense has been extended and commented on by several authors (see for example Baroni, 2007) in the field of literary theory. Sternberg's discussion of these concepts give some additional insight into Brewer and Lichtenstein's work. In his 'Narrativity: From Objectivist to Functional Paradigm' (Sternberg, 2010, p.640), Sternberg presents his three master elements of narrative dynamics:

- prospection,

- retrospection,
- recognition,

or in Brewer and Lichtenstein's terms: suspense, curiosity, and surprise. Sternberg grounds these dynamics in the 'ongoing survival value of observing, plotting, telling, foretelling, inferring event lines' (ibid., p. 607), and also suggests that these different narrative dynamics occur separately from each other, so that:

> 'the *prospector* looks ahead to some contingency and the *retrospector/recognizer* looks backward on some mystery, with a view to closing gaps opened on the move between them.' (ibid., p. 640)

In his reading, suspense thus depends on the prospective mind and 'arises from rival scenarios envisaged about the future'. This prospective mind is: 'uncertain (e.g., both hopeful and fearful) regarding the outcome suspended and restlessly shuttling between the imagined (e.g., hopeful/ fearful) outcomes.'

Curiosity is similar to suspense, but directed towards the *past* and not the future:

> 'the curiosity-driven processor expects ultimate stable closure of the fragmentary, disorderly data, but meanwhile needs to supply it as best one can when left under-informed, via tentative, multiple, often incompatible, always revisable gap-filling hypotheses.' (ibid., p. 641)

As for surprise, Sternberg says that 'we must be lured into false certainty for a time about time past', and then:

> 'a hypothesis established beyond doubt, (fact-like in our eyes, rather than uneasily forked, as in prospection and retrospection) will collapse with a vengeance and give place to some other...'

The universality of suspense, curiosity and surprise Sternberg also emphasises the universality of these three narrative dynamics:

> '...Everything in narrative must accordingly go through the twin process of happening-cum-telling/reading hence through the dynamics of suspense, curiosity, surprise and influence it in turn...' (Sternberg, 2009, p. 501)

Thus, everything in a narrative is treated in terms of suspense, curiosity and surprise, based on the *Syuzhet/Fabula* distinction. Secondly, these three dynamics can map themselves onto any surface form.

> '...We thus map suspense (i.e., our felt uncertainty about the narrated future) onto an impending conflict, or the narrator's wink ahead, or the hero's fear, or a proleptic epithet, or a traditional happy/unhappy closure in doubt, for example;...'

Furthermore, he claims that these three dynamics can 'narrativise' almost any element of a given discourse:

> '...Even components and structures that narrative shares with nonnarrative texts or with textuality at large, such as spatialities, characters, viewpoints, themes, ideology, semiotic code (e.g., language), and time of communication itself, assume a distinctive reference and energy once controlled and mobilized by the dynamics of narrative...'

Sternberg further suggests definitions of *narrativity* and *narrative* based on these elements:

- Narrativity: 'the play of suspense/curiosity/surprise',
- Narrative: 'a discourse where such play dominates'.

and claims that these definitions capture 'both the genre's immense variety and our intuitive knowledge of its unity as no other definition has.' The universality that Sternberg has claimed for suspense, curiosity and surprise underlines the need for good theoretical models of these phenomena if we are going to acquire a deep understanding of narrative.

2.4 Computational models of narrative

2.4.1 Introduction

We will first review a representative sample of computational approaches to *narrative*, examining their relevancy to the development of a model of suspense. We ask the following questions of each approach:

- How *would* or *could* this model generate a suspenseful story? Explicitly or as an emergent characteristic of the story?
- What concepts would the generation of suspenseful stories *rely on* in this model?
- Which components of the model shed some light on how suspense *could* be modelled?

Following Bailey (1999), we group models of narrative into the following categories:

- Character models (also called Autonomous agent models and World models)
- Story models
- Author models
- Reader models
- Interactive narrative models

To illustrate each different type of modelling, we will examine one or two examples for each category.

We then examine some computational models of *suspense*, before discussing some *knowledge acquisition systems* for generating narrative inferences from the real world, and describing some different *suspense typologies*.

2.4.2 Character models

Character-based computational models of narrative were some of the first systems to be developed. Such models build up a representation of a storyworld which behaves according to a set of rules and which contains a number of autonomous characters. Storyworld modelling is in general based on some explicit causal and/or intentional structure. Most character models use the concept of character *goals*. Characters generate and carry out actions in order to realise their goals. The story emerges from the interactions of the characters as they attempt to achieve their goals in the storyworld.

One of the most well-known of such systems is the TALE-SPIN system (Meehan, 1977, 1981), which creates stories by simulating a forest storyworld, assigning goals to characters in this storyworld and defining what happens when these goals are pursued.

In the VIRTUAL STORYTELLER (Theune et al., 2003), just as in TALE-SPIN, stories emerge from the actions of the characters. The range of events and states used is, however, much larger than in the TALE-SPIN system and includes for example, characters' cognitive states such as beliefs and emotions, as well as characters' goals, actions and perceptions. In this system, the story plot is constrained by a director module that can disallow actions by characters if they would conflict with the construction of a well-structured plot. The plot concept that is used is based on Freytag's Triangle (see 2.2.2).

We now examine the following aspects of these two character models:

- Their inferential mechanisms,
- Their premises for cognitive and emotional interest,
- Their potential for suspense modelling.

Inferential mechanisms

For character models, the goal is a high degree of realism in the modelling of the narrative world and its characters. Inferential mechanisms of *character* models are of course principally based on the causality of the 'properties' of *characters*. The characters are thus modelled as entities that have certain *psychological* properties and capabilities. Causality is usually dualistically separated into physical and psychological categories and the definition of the storyworld and its events prescribes the *physical* causal inferences that can be made.

The VIRTUAL STORYTELLER (Theune et al., 2003) creates causal connections between story elements, thus fixing the background network of causality of the storyworld. If for example, some algorithm determines that a certain character goal G motivates an action H to achieve the goal, this is stored as a causal link $G \rightarrow H$. Once these causal links have been created, they are linked together in a causal network, which is a representation of the *Fabula* of the storyworld. From this *Fabula*, events can then be chosen to create a *Syuzhet*, or a particular telling of the story.

Premises for cognitive and emotional interest

TALESPIN (Meehan, 1977, 1981) allows a description of the interactions between its characters. They can be hostile, friendly, honest or dishonest with each other. In the VIRTUAL STORYTELLER (Theune et al., 2003), characters in the storyworld have a more sophisticated mental and emotional model to help them choose which actions to undertake. Nevertheless, like all character models, their success is based on the following premises:

1. Emotional interest:

   i) Autonomy $\Rightarrow$ Believability. Autonomous characters will be able to maintain reader interest by being believable.

   ii) Centrality $\Rightarrow$ Empathy. Using characters as the central element in story construction must create reader empathy and emotion.

2. Cognitive interest:

   i) Goal conflict $\Rightarrow$ Interest. If the characters have some interesting conflicting moments in the storyworld, then this will be enough to create a story.

Potential for suspense modelling

Meehan's theory of narrative as implemented in TALE-SPIN is simple: 'a story is about a problem and how it gets solved'. If we allow the assumption that a problem in a character-based world must always be based on some kind of conflict, then TALE-SPIN's link with the notion of suspense could be described thus:

$$\text{Problem} \Rightarrow \text{Unresolved conflict} \Rightarrow \text{Suspense}$$

There is however, in character models, a strong separation between the storyworld modelling and the effects that the story produces. Such models can achieve a good level of character believability, but are not so good at creating coherent plots. Furthermore, they neither guarantee that conflict will arise, nor that any conflict will be interesting or last long enough to generate suspense. Indeed, TALE-SPIN's most important lesson is that when stories are only driven by character goals, uninteresting stories often get created.

To summarise: in character models, suspense is usually not controlled or explicitly modelled and any suspense that does result comes by chance from the conflicting situations that the characters find themselves in.

2.4.3 Story models

Story models use some different kinds of narrative representation as their starting point. We will examine here two types:

- Story grammar models
- Plot-based models

Story grammars

In story grammars, events in a narrative are interpreted as being cases of a *type* of narrative component, much in the same way that a word belongs to a syntactic category. The development of story grammars can be seen as an attempt to render declarative narrative models amenable to computational implementation. Story grammars can then be linked to the generation processes used by human authors.

One of the first story grammars was developed by Rumelhart (1975) and was designed as part of a theory of story summarisation. It uses the following syntactic rules for the creation of a well-formed story:

1. Story $\Rightarrow$ Setting + Episode
2. Setting $\Rightarrow$ (State)*
3. Episode $\Rightarrow$ Event + Reaction
4. Event $\Rightarrow$ {Episode | Change-of-state | Action | Event + Event}
5. Reaction $\Rightarrow$ Internal Response + Overt Response
6. Internal Response $\Rightarrow$ {Emotion | Desire}
7. Overt Response $\Rightarrow$ {Action | (Attempt)*}
8. Attempt $\Rightarrow$ Plan + Application
9. Application $\Rightarrow$ (Preaction)* + Action + Consequence
10. Preaction $\Rightarrow$ Subgoal + (Attempt)*
11. Consequence $\Rightarrow$ {Reaction | Event}

These rules can be seen to combine elements from different domains:

Typical story features: Episode, Setting, Event

Basic psychological events: Internal Response, Emotion, Desire

A simple planning grammar: Rules 8-11

The JOSEPH story generation (Lang, 1999) is the first such system to be constructed from an explicit, formal model for stories and uses a story grammar similar to that of Rumelhart. Stories have two components, a setting and a list of episodes. In addition, each episode has the following four parts:

1. An **initiating event**
2. An **emotional** response on the part of the protagonist
3. An **action** response on the part of the protagonist
4. An **outcome** or state description which holds at the conclusion of the episode.

This story grammar is similar to that of Rumelhart but with the planning elements removed. We can in fact derive it from Rumelhart's grammar by the following rewriting steps:

- Episode ⇒ Event + Reaction
- Episode ⇒ Event + (Internal Response + Overt Response)
- Episode ⇒ Event + {Emotion | Desire} + {Action | (Attempt)*}
- Episode ⇒ Event + {Emotion | Desire} + {Action | (Plan + Application)*}
- Episode ⇒ Event + {Emotion | Desire} + {Action | (Plan + (Preaction)* + Action + Consequence)*}

By simplifying and regrouping this last form, we can achieve a form similar to that used by Lang:

- Episode ⇒ (initiating) Event + Emotion + Action + Consequence

We can see here that this definition of an episode contains both internal (or psychological) and external (or physical) causalities.

In the BRUTUS top-down story generation system (Bringsjord and Ferrucci, 2000), the starting point is always a literary theme in the form of high-level story schema such as 'betrayal', that is chosen for its intrinsic interestingness. The schema is worked on by a 'world-simulator' which combines a storyworld model with logic and causality rules to produce a kind of instantiated thematic plot. There then follows a hierarchy of paragraph and sentence grammars which produce the final textual form of the story. BRUTUS thus uses grammar-like techniques from the very highest level right down to sentence structure.

Causality and Story grammars Black and Bower (1980) have criticised story grammar models for their lack of rigour, claiming that they are incapable of distinguishing between stories and 'non-stories'. They proposed a theory based on state transitions in causal chains of events. However, their criticisms have also been questioned by Mandler and Johnson (1980). Work on the representation of narrative in memory by Trabasso and Van Den Broek (1985) suggested that causality is more important than story grammar inclusion for event representation in memory, but there is debate about the relative importance of story grammars and causality in narrative generation and understanding.

Premises for cognitive and emotional interest The story grammar premise is that if a narrative is in some sense well-formed, then it will be successful, that is, entertaining. However, the story grammar in Rumelhart

(1975) includes Emotion and Desire as Internal Reaction to Events and this inclusion can perhaps be seen as an additional mechanism to ensure a degree of *empathetic interest* over and above the well-formed nature of a story. Similarly, the goal-oriented structure of characters' actions is there to provide a minimum level of cognitive interest.

BRUTUS (Bringsjord and Ferrucci, 2000) offers perhaps the best guarantee for cognitive and emotional interest because it uses themes that are chosen *from the outset* as intrinsically interesting. This design choice is in itself of interest; there would seem to be a range of themes similar to betrayal (jealousy, revenge, overcoming hardship, etc.), that contain just the right dosage of emotional and cognitive complexity for human readers to be used as the basis for interesting and even compelling stories. The existence of a theme-level in narrative is an interesting question for future research.

Potential for suspense modelling As we have seen, story grammars have a basic structure leading from an initiating event through actions and emotions to a final outcome event. This structure is of course, also used by the model of suspense in Brewer and Lichtenstein (1982), and it appears to be paradigmatic for suspense.

In BRUTUS (Bringsjord and Ferrucci, 2000), the use of a structured theme like 'betrayal' allows the system to author a series of events which *could* (or one might say *should*) elicit suspense: 'will my 'friend' *really* give me the money he promised. . . ' Our concern is however, just *how* do such narrative structures achieve this? There are many such themes possible, but how is the suspensefulness of their corresponding stories generated and maintained?

Plot

Other story models explicitly use the concept of *plot*. The role of the characters and their motivations in a story is central to the elaboration of plots. There is no over-arching structure such as for example 'betrayal' to guide the sequence of events. Aesthetically, the role of plots in a story can be more satisfying when they are 'total' in the Aristotelian sense, that is, when in some sense the plot explains and motivates every action in the story. Chatman (1980) proposes a distinction between *kernels* and *satellites* in this respect. *Kernels* are events that move the plot forward, 'by raising and satisfying questions' and *satellites* are less important events which can be left out in the telling of the story without disturbing the logic of the plot.

Lehnert (1981, 1982) has criticised story grammars for not being general enough to capture very different variations in plot structure. In Lehnert's theory of plot units there is no pre-determined over-arching structure, it is rather the affective states of the characters that build a plot. Characters' affect-states come in three types in this simplified model: positive states, negative states and mental states with neutral emotionality. Affect-states are causally linked to other affect-states and events in the following different ways:

$$\text{Event} \overset{\textit{motivates}}{\Longrightarrow} \text{Affect-state}$$

$$\text{Affect-state} \overset{\textit{actualises}}{\Longrightarrow} \text{Event}$$

$$\text{Affect-state} \overset{\textit{motivates}}{\Longrightarrow} \text{Affect-state}$$

$$\text{Affect-state} \overset{\textit{terminates}}{\Longrightarrow} \text{Affect-state}$$

By combining rules like these, quite complex plot units can be constructed to characterise plots such as 'Problem solution by effective coercion' or 'Double-cross'. Lehnert's work has recently been revived by Goyal et al. (2013) in a system called AESOP which automatically generates plot unit representations for narrative texts by using four steps: affect state recognition, character recognition, affect state projection and link creation.

In a similar way, the PLOT MANAGER developed by Sgouros (1999) calculates possible behaviours for each character and then tries to combine these behaviours into interesting sequences. Four different types of dramatic situation are used:

- Lifeline: a character has a chance to improve their situation
- Rising complication: a bad situation gets worse
- Reversal of fortune: a good situation turns bad
- Dramatic irony: the interaction between two characters is not reciprocal in a kind of story twist.

An important characteristic of this approach is that characters have to overcome some difficulty to fulfil their goals, possibly in the form of a personal conflict.

Premise Plot-based story models are based on the premise that the character-centred interplay of motivation, action, reaction and event will be interesting in itself, provided that there are *enough* conflicting moments.

Potential for modelling suspense Plot models have in common the concept of conflict; because characters act in a storyworld and attempt to overcome problems and conflicts, drama can arise. Sgouros's dramatic situations have emotional importance built-in. They can also be seen as containing an initiating event which signals a potentially successful process with a clear outcome, and in this they have a clear potential for suspense evocation.

2.4.4 Author models

Author models attempt to model the way a human author goes about the task of creating a story. They often have a top-down approach, but are different from story grammar models in that there is an explicit role for the author to decide on the form and content of the narrative. Many author models combine aspects from story grammar models, characters models or even reader models. We concentrate here specifically on how involvement of the author can give insight to suspense in narrative, again examining the following aspects:

- inferential mechanisms
- premises for cognitive and emotional interest
- potential for suspense modelling

Inserting story elements

Lebowitz (1985) developed UNIVERSE, a model of story telling based upon an extensible library of plot fragments. Plot fragments resemble a series of writer's aids for the creation of story-telling universes and contain characters and their histories, family relations and interpersonal relationships. The system creates plot fragments using an algorithm driven by *author goals* rather than character goals. For example, an author can have the goal to keep two lovers apart, and thus insert story elements that stop them from meeting.

Obstacles such as these can of course increase the dramatic interest of a story. The approach fits with the idea that characters' actions are ultimately motivated by the author's goals in telling the story and not the characters' goals inside the story.

Barber and Kudenko (2008) created an interactive narrative system called GADIN which attempts to create dramatic tension in a similar way by using dilemmas such as 'Betrayal', 'Sacrifice' and 'Greater Good' which users must overcome. The dilemmas are inserted into the ongoing interactive plot to increase the tension. The working premise is that dilemmas create conflict which builds tension which can produce dramatic interest.

Clearly, this procedure can create or maintain certain conflictual situations and play a role in increasing suspense. However, there is no actual model of suspense in these approaches. There is only a series of methods for manipulating a narrative which may or may not actually affect perceived suspense.

Balancing author and character goals

MINSTREL (Turner, 1993) is a computer program that uses both authorial and character goals. It generates short stories about King Arthur and his Knights of the Round Table. The system uses case-based reasoning to treat story generation as problem-solving and uses four types of hierarchically linked authorial goals:

- Thematic goals
- Consistency goals
- Drama goals
- Presentation goals

In MINSTREL, all the elements that comprise a story are represented as schemas. MINSTREL's main contribution is the concept of transform-recall-adapt methods (or TRAMs), which demonstrate the creative power of small changes in story schemas.

Ware and Young (2014) produced a state-space narrative planner called GLAIVE which also attempts to integrate both author and character-based approaches. It creates stories that are clearly motivated and goal-oriented for the characters in them and which at the same time satisfy the author's narrative goals. One overall plan both represents the entire story and contains sequences of steps which correspond to the characters' plans. These in turn are described in terms of characters' goals and causal structures.

Causality is defined by causal links and intentional paths. We can think of this as physical and psychological causality. Furthermore, all causal steps should be able to be 'explained'. A step is taken as being explained if it is part of a character's plan even if that plan fails. Steps can have causal 'parents' and causal 'ancestors'. GLAIVE explicitly defines how earlier steps satisfy the preconditions of later steps:

- every step in a series of causal links has a causal parent and causal ancestors, and
- every step in a character's intentional path must be intended to be true by the character.

Both MINSTREL and GLAIVE attempt to balance author and character goals during the development of a narrative. However, neither offers a model of what makes a narrative suspenseful. The modelling of causality in GLAIVE is nevertheless interesting for our purposes because it assumes that physical and psychological types of inference have equal importance in story comprehension. This is a reminder that it is not the *type* of inference that is important for the modelling of narrative phenomena and this insight is important in the development of our model.

Tension as a parameter

MEXICA (Pérez and Sharples, 2001, Pérez y Pérez, 2007) is an author model of narrative generation based on a cognitive account of writing which attempts to create novel and interesting stories. A story is deemed to be interesting when the tension in the story varies due to the variation in the characters' well-being. MEXICA calculates tension from evaluations of the variables *love*, *emotion* and *danger* which are based on links between the characters. The tension is represented numerically at all moments in the story. The system then compares the tension of the current story with that of previous stories to evaluate its interestingness. In MEXICA a story is defined as a sequence of Linguistic Representations of Actions (LIRAs), and the system requires a dictionary of LIRAs to work. LIRAs are actions that characters can perform in the story whose consequences change the storyworld in some way. Each action has a set of pre- and post-conditions which can be of two types:

- emotional links between characters
- dramatic tensions in the story

Thus, in MEXICA, two types of causality are used and the only explicitly causal modelling of the storyworld itself is ensured by the emotional links between characters. The system thus has the *explicit* goal of creating stories with emotional interest to engage the reader.

The interactive drama system FAÇADE (Mateas and Stern, 2003) also uses tension in the form of an ideal tension curve which serves as a guide for the development of an interactive narrative. The groups of events or 'beats' in the story can be presented in a number of ways while still containing more or less the same information. The system chooses a way to present the event in order to obtain the desired tension level at that point in the story. The tension parameter functions essentially as a guide to decide how best to select and present the next event so that the story follows an Aristotelian dramatic arc (see 2.2.2).

The interactive fiction system by Barros and Musse (2008) also uses curves to represent dramatic tension. The system has a definition of narrative tension based on the discovery of *clues* by the player. The narrative generation is again guided by trying to find the best-fit between the actual and ideal tension curves.

The use of tension in the above systems is an indicator of their potential to successfully model suspense. However, even though tension is often linked to suspense, none of these systems are trying to explicitly make the story more suspenseful. Rather they attempt to create the perfect dramatic arc. We surmise however, that the perfect dramatic arc is the *result* of suspense mechanisms rather than the other way around.

2.4.5 Reader models

Reader models are based on modelling the response that a story creates in a reader. Of course, reader models have the premise that, by including the reader's reactions in the development of the narrative, they will be able to vary the narrative's level of interest.

Bailey's story generation system (Bailey, 1999) is based on a model of the responses of a typical reader. The story generator is guided by a heuristic that seeks to achieve optimal 'storiness' which is defined in terms of the *expectations* and *questions* generated by the reader. Questions are also considered more important than expectations, as sequences of expectations without obstacles, that is, without a level of uncertainty, would be uninteresting.

Szilas' interactive narrative architecture IDTENSION (Szilas, 2003) includes a model of the user which attempts to estimate the effect of each possible action on the user. To do this, it uses the following narrative criteria: character motivation, character ethical consistency, relevance to previous actions, and conflict. At any given step in the story, first, all possible actions are generated. These actions are then evaluated and filtered according to how the user of the interactive narrative would perceive them. Szilas concludes the presentation of IDTENSION by saying that the stories produced lacked 'dramatic intensity', and deduces that merely increasing emotional involvement is not sufficient to create a strong narrative.

For reader models, the modelling of the reader's reactions alone does not seem to be enough to derive or suggest specific ways to model the storyworld. As a result, the storyworld modelling in such models seems somewhat arbitrary. However, because in such models, the narrative modelling is concerned directly with the reader's reactions, such models are perhaps the closest to being able to model and generate suspense in stories.

Interactive narratives

In all narrative an important additional question is, of course, to what extent the reader feels *immersed* in the narrative by identifying themselves with the characters in the story.

If this immersion occurs in an interactive narrative situation, then the interactive narrative resembles the character models because the reader can interact in the story just like one of the story characters. The character model premise might then apply: "if the characters (in this case, also the interactive reader) in a storyworld pursue their goals, then an interesting story will result".

In this respect, the affect-detection module developed by Zhang et al. (2008) attempts to make inferences about the affective states of human-controlled characters in an improvised e-drama system, by analysing the characters' textual speeches. The goal of the module was to enable the partial automation of a director character that could intervene in the improvised drama sequences. Feedback about the played characters' affects can, however, also be seen as a step towards determining the degree of *immersion* of participants in interactive narrative systems as they control their respective characters.

Furthermore, depending on the nature of the interactive system, participants might be able to change the flow of the narrative and thus act somewhat like authors. However, in general, the author of the system will still know much more about the storyworld, the story modelling and the various constraints on the narrative that are present in the system. The interactive participant usually makes choices *about which* the author has decided. Interactive models can thus be seen as a mix of author, reader and character models.

2.4.6 Computational models of suspense

None of these systems thus far give an explicit formal analysis of how suspense in narrative could be generated. The focus is mostly on the *global* story-modelling task and on the automatic generation of new narratives. We now look at computational models of narrative which are explicitly constructed around the concept of suspense.

Since 1995, the Liquid Narrative Group at North Carolina State University has developed interactive narrative approaches based on planning. Several approaches explicitly use suspense in their story generation processes. We briefly review a selection of this work and then describe DRAMATIS, a related approach.

Suspense through outcome management

Cheong and Young (2006) describes a planning-based approach which models the goals and actions of characters in a storyworld and attempts to specifically design and generate narratives that evoke suspense. It uses a definition of suspense taken from Gerrig and Bernardo (1994), which claims that the suspense level readers feel depends on the number and type of solutions they can imagine in order to solve the problems facing the main protagonist: "the reader's suspense is heightened when undesirable outcomes are likely to happen over preferred outcomes" (ibid. p.2). The focus of the system is on the suspense created uniquely by the *story structure*.

Suspense through event insertion

Cheong and Young (2008) proposes SUSPENSER which attempts to find a suspenseful telling of an existing story by simulating the reasoning process of an implied reader using a planning algorithm. It is one module in the following sequence of modules which is labelled with the terminology of the Russian narratologists (see 2.2.3):

- a Fabula Generator which creates a Fabula

- a SUSPENSER which turns the Fabula into a Syuzhet
- a Discourse Generator which turns the Syuzhet into a medium

Once the basic storyline or Fabula has been established in the form of causal connections between a series of events, SUSPENSER attempts to find events which could be *added* to the story to increase the suspense level. SUSPENSER is also based on Gerrig and Bernardo's definition of suspense (Gerrig and Bernardo, 1994); the reader should feel more suspense when the number of possible ways for a protagonist to escape are reduced.

Surprise linked to curiosity

Bae and Young (2008) proposes PREVOYANT which works with surprise in a similar way to SUSPENSER with suspense[2]. The system uses the model of surprise proposed by Brewer and Lichtenstein (1982) which uses the concept of an Initiating Event. To create surprise, an event is revealed without part of the causal chain leading up to it, that is, without some of its initiating events. To do this, the system changes the order of events in the *Fabula* to a non-chronological one, creating a *Syuzhet* that uses flashback and foreshadowing effects to create surprise. The revelation or inference of the existence of the missing events then resolves the curiosity that the surprising moment created.

Linking suspense and surprise

Similar work by Bae and Young (2009) using a plot model, explores the *relationship* between suspense and surprise. This model uses the concept of

[2]Although the system does not deal directly with suspense, we include a brief description here for completeness, and also because our research goal is to find a model of suspense within the framework laid out by Brewer and Lichtenstein which is based on suspense, curiosity and surprise.

'disparity of knowledge': for suspense, the reader often knows *more* about the story than do the characters, whereas for surprise the reader often knows *less*.

The model makes a distinction between the plot the reader currently believes is true of a given story from another more accurate one, known to one of the characters. At a certain moment in the story, an event occurs which forces the reader to change their reading of the story, thus creating surprise. This approach again follows the definition of surprise given by Brewer and Lichtenstein (1982), where part of a causal chain of events is first hidden, and then suddenly revealed.

Suspense with foregrounding

O'Neill and Riedl (2014) proposes DRAMATIS, which is also based on the definition of suspense in Gerrig and Bernardo (1994). The two major components of the system are the following:

- an algorithm which tries to determine the most likely escape plan for the main protagonist, as perceived by the reader.
- a model of *reader salience* which attempts to model which narrative events are the most foregrounded in the reader's mind at any time in the telling of the story.

Conclusions

Plans and conflict Cheong and Young (2008) provides some interesting feedback on the SUSPENSER system, describing the difficulties involved in combining a planning paradigm with suspense modelling. Plans are usually considered good solutions to a problem in a situation where there are no conflicts. But one of the basic requirements for suspense is that conflict is present. They state that 'protagonist's and antagonist's plans were often related via causal relationships'. They express the need for the development of a 'more conflict-expressive plan representation' for suspense modelling.

As we shall, our storyworld modelling technique takes into account the necessity for different characters' plans to be causally linked.

Gerrig and Bernado's definition of suspense Another difficulty that the authors of SUSPENSER mention is the relationship between Gerrig and Bernado's suspense definition and their plan model. The definition of suspense given in Gerrig and Bernardo (1994) states that readers feel more suspense *when the number of possible ways for a protagonist to escape is reduced*. By attempting to determine the ratio of failed plans to successful plans as a way to capture this definition, they encountered the difficulty of determining just what counts as a failed plan. Failed plans can occur for many reasons in a plan model and not just because they fail in the storyworld. Attempting to fit this definition into a plan model therefore appears somewhat *ad hoc*, and, it seems, also poses certain computational difficulties.

Even though Gerrig and Bernado's definition of suspense has some support in the psychology literature, we wonder whether it is a rather too specific description of suspense, rather like a special case. Of course, many typical suspense stories do use an increasing threat of danger to the main character to generate suspense, and the definition used in Gerrig and Bernardo (1994) may be a useful description of such cases.

We will argue, however, that the link between increasing suspense and a reduction of the number of escape routes for a protagonist is due to the parallel reduction in the *ambiguity* or *uncertainty* of the situation that such a reduction produces. The resulting increased *confidence* in the escape routes that are left boosts the simple *conflict-based* suspense that is present between these escape routes and the dangerous outcome.

We discuss this further in 3.4 and show how such situations can be modelled using the two *different* types of suspense mechanism present in our approach.

Reader salience DRAMATIS (O'Neill and Riedl, 2014) is the first computational model of suspense we have encountered that uses a model of reader salience, or foregrounding. As suspense is a dynamic phenomenon, one might expect the suspense level to fluctuate during the telling of a story; a certain set of causal links or goal paths may have less of an effect on the reader when they are not mentioned for a while. DRAMATIS attempts to show a way towards finer grained models of narrative which are capable of handling complex multi-threaded stories.

Our storyworld model will also propose ways to take into account both suspense fluctuations and multiple story threads.

2.4.7 Computational models of narrative inference

We now discuss some techniques to computationally model and source inferential processes for narrative comprehension and generation.

Causal networks and goal hierarchies

The planning or goal-based approach to story modelling is used in many systems as we have seen. In some ways, plan models can be seen as an extension and improvement of the linear concept of a *Fabula*. However, planning models tend to be oriented towards the use of one single final goal in the future. In contrast, stories contain many open-ended events which can have multiple effects on future events for different characters. A degree of forward-branching seems essential to capture this, and this is not the habitual function of plans. There is further evidence which seems to downplay the importance of planning structures in narrative comprehension and we will now discuss this briefly.

Trabasso and Van Den Broek (1985) tested the relative importance of goal-based hierarchies of events and causal event networks in human representations of stories. They came to the following conclusions:

- A goal statement's change from a superordinate to a subordinate level decreased its probability of being included in a summary *only if* this shift was accompanied by a change in its causal role. When the number of causal connections and the causal chain status were held constant, the hierarchical level had no effect.
- When the number of causal connections increased, the likelihood of summarization for both goal and other statements increased.
- Causal relations are operative and transitive, in that the strength of the relations declined linearly as a function of causal distance in the network representation, independent of temporal and reference distance.

They summarise these results by suggesting that 'the importance of a statement in a structure is the result of causal reasoning during comprehension.' They also suggest that there is a natural hierarchy of importance for the different types of causal connections, going from the most important to the least important: 'physical, motivational, psychological, and enabling relations.' These findings suggest a relatively weaker role for goal hierarchies in humans compared to that played by causal network representations.

Interestingly, the new search planning algorithm proposed by Riedl and Young (2010): the Intent-based Partial Order Causal Link (IPOCL) planner, includes more causality than previous versions. It also 'reasons about character intentionality by identifying possible character goals that explain their actions and creating plan structures that explain why those characters commit to their goals.' The authors' evaluation of this planner shows that it generates plans with improved audience comprehension of character intentions compared to other partial-order planners.

The relative importance of planning algorithms and causal links in narrative modelling is a direction for further research.

Sources of inference from the real world

Narrative systems often need some kind of background knowledge. Different proposals to automate the acquisition of this knowledge have been put forward. It seems that the knowledge representation bottleneck which is so problematic for so many approaches to narrative may be starting to be overcome. We now briefly review some approaches to knowledge acquisition which could be useful for narrative systems.

PHARAOH (Hodhod et al., 2012) is a context-based structural retrieval algorithm for cognitive scripts that uses keywords and semantic structures. Notably, by considering the timing of events in a script, the system allows the retrieval of cognitive scripts according not only to their structure, but also to their context.

The MAKEBELIEVE story generation method (Liu and Singh, 2002) uses common sense knowledge transposed into frames together with 'fuzzy, creativity-driven inference' to generate short fictional texts based on the first line of the story which is supplied by the user. Here is an example:

> John became very lazy at work. John lost his job. John decided to get drunk. He started to commit crimes. John went to prison. He experienced bruises. John cried. He looked at himself differently.

Boujarwah et al. (2012) attempts to delegate the acquisition and aggregation of procedural knowledge to *large collections of people* rather than to automated processes. Participants were asked to list typical actions for a given context such as a restaurant, then for each action, to make a list of potential obstacles and also possible solutions to these obstacles. Their responses were grouped together by finding action synonyms. A graph of the acquired knowledge was then made which combined all the steps, obstacles and solutions.

O'Neill et al. (2014) proposes a formalised coding procedure inspired by qualitative research methods to create narrative knowledge from a corpus. Coders identify common actions and themes in a corpus over a number of iterations. These actions and themes then are used as a code taxonomy that can be used by many coders to generate knowledge structures which correspond to the specific representations that are needed for a given system. The authors also describe how this method was used in the context of the narrative system DRAMATIS (O'Neill and Riedl, 2014).

Finally, Chambers and Jurafsky (2009, 2010) describes the extraction of narrative schemas called 'narrative event chains' from newspaper articles using a machine learning approach. These event chains are a kind of script which describes what kind of events typically follow each other and what active or passive roles are filled by the actors involved. This unsupervised system aims to provide varied and rich inferential structures that can be used by other narrative systems. Indeed, some aspects of our model are specifically designed to interface smoothly with the event chains that the authors describe.

2.4.8 Suspense typologies

In O'Neill (2013), which extends the work started in DRAMATIS, some different types of suspense are distinguished:

- *Procedural expectation suspense*
- *Outcome expectation suspense* of which *surprising suspense* is a subset

Procedural expectation suspense is based on viewers knowing what upcoming events or obstacles are likely to occur in a story. O'Neill describes two variants of this type:

- *genre-knowledge suspense* where the information about upcoming events comes from knowledge about a given narrative genre, and
- *opposition suspense* where the knowledge about events comes from information in the story-so-far about the characters plans and goals for example.

Outcome expectation suspense is according to O'Neill, a different kind of suspense where viewers desire a certain outcome but have no idea about how it could come about. O'Neill goes on to say that his current model, which is based on Gerrig and Bernardo's definition, cannot deal with outcome expectation suspense.

The model we will propose abstracts away from the source of the information available to the viewers and attempts to concentrate on the *structure* of the information flow. In our terminology we will describe **conflict-based suspense** which corresponds roughly to O'Neill's procedural expectation suspense. In our opinion, the only difference between the two sub-types that O'Neill proposes: genre-knowledge suspense and opposition suspense, lies in their source of information and we question the utility of this distinction for a general suspense model. We will also provide as an integral part of our system, a first formal model of **revelatory suspense** which corresponds roughly to O'Neill's outcome expectation suspense. In our terms, O'Neill's surprising suspense can be seen as an extreme type of expectation suspense where viewers have *no idea at all* about what might occur. In such situations there will be a surprise whatever happens.

2.4.9 Summary

In a similar way to pre-computational models of narrative that concentrated on the notion of plot, in many computational models of narrative, a frequent approach is to determine some basic element, which, when manipulated in certain ways, will produce a skeletal story-line at the plot level.

As we have seen, systems such as Meehan's TALE-SPIN (Meehan, 1977) uses the characters' *goals*. MINSTREL (Turner, 1993) uses *both* authorial and character goals. MEXICA (Pérez and Sharples, 2001, Pérez y Pérez, 2007) uses a *tension curve* to represent *love*, *emotion* and *danger* in order to drive the generation process. Cheong and Young (2006) uses a *planning*-based structure which models the goals and actions of a series of characters who belong to the given storyworld.

Each system attempts to use the storyworld structure they put forward to generate stories that are in some sense interesting or plausible and thereby suggest the validity of their approach. The adequacy of narrative generation is thus often seen as the litmus test of any computational theory of narrative, and many systems focus on this aspect.

Apart from the reader models, many systems can also be seen to follow another tendency of pre-computational approaches in that they do not make explicit a *theory of narrative comprehension*.

Further, none of these systems give an explicit formal analysis of how suspense is created. The focus is rather on the *global* story-modelling task and on the automatic generation of new narratives. Suspense is often seen as just one of a set of by-products of story generation which must be present for a story to be interesting and the goal of creating suspenseful stories is often baked-in to each individual system. There is no portable idea of what makes a suspenseful story.

It is our view that by focussing only on the global story-telling task, such systems may suffer from a degree of arbitrariness in the choice of their theoretical story modelling apparatus. We believe that the presence of more systematic and fundamental approaches to suspense (and to other aspects of what makes a story entertaining such as curiosity and surprise), could help to create a common ground for the evaluation of story modelling systems.

2.5 Conclusions

In the preceding chapters, we have reviewed a series of models of narrative from a wide range of fields, examining their differing relationships to suspense in narrative. The approaches to the question of narrative in general and suspense in particular are very diverse. Cheong and Young (2006), as we have seen, created a heuristic for suspense based on narrative modelling using the planning paradigm. Characters have goals and corresponding plans, and suspense levels are calculated as a function of these.

However, the presence of suspense is *not* dependent on the existence of characters' goals: we can experience suspense about a ball rolling off a table, or a piece of string breaking under the strain of a weight, or an ice floe breaking up. Indeed, the mere *existence* of such types of suspense has been one of our motivations in the search for a more portable domain-independent model of suspense in narrative.

Many approaches to modelling narrative concentrate from the outset on a complete model of the whole phenomenon. The starting point is the 'big picture' about which we all believe that we know something: how to tell a story. However, this 'big picture' may be simply too wide-ranging for our current knowledge. We suggest that our attempts to model narrative are perhaps rather like making a model of how the body functions with only the vaguest intuitions about its different internal organs, or like a model of vision that neglects the fact that objects can be placed at different distances. Approaches to narrative can then fall into the same trap as the how-to-write-a-story literature; much is assumed, but each proposal assumes *slightly different elements*.

There is a clear lack of consensus on theoretical approaches to suspense. some of the approaches we have covered are formal and prescriptive but lack independence from other concepts such as plans, others approaches give independent but rather general and descriptive accounts of this phenomenon.

One of the goals of this research will be to contribute to a precise formal definition of suspense which could perhaps be applied, amongst other things, to the planning narrative paradigm, but which would not *depend* on such a context. Our working assumption is that:

- at least one underlying model of suspense exists and can be applied to a wide range of domains and contexts.

We are aiming for a independently motivated *domain-independent* account of what makes a suspenseful story.

Chapter 3

Towards a domain-independent model of suspense

Our research goal is to propose a model of suspense phenomena in narrative that is as domain-independent as we can make it. In this chapter, we build up and discuss in an informal way a number of the elements that we think such a model of suspense should include. This will prepare the ground for a formal, mathematical treatment of our model in the following chapter (4).

3.1 Suspense in the real world: a sliding puck

To better identify some of the aspects of suspenseful situations that we will be using to build up our model, we describe a thought-experiment around the sport of curling.

We imagine a standard situation where a puck has been pushed forwards by one of the players and is sliding towards its target across the ice. We will focus on the suspense that can be felt once the puck has been released, in that quite long period lasting several seconds as the puck approaches the target stone and slowly comes to a halt[1].

This situation was one of the simplest suspenseful situations we could find. It also has the advantage of being suspenseful even though there is no character acting according to internal plans or goals. It will thus serve as a safeguard in our suspense modelling, helping us to avoid the use of in-built anthropomorphic features. We can then compare the features present for the sliding puck with the equivalent features for characters in a story.

Our first question is:

- While watching the puck of our favourite team slide nearer and nearer the target in curling, what are we doing?

We suggest we are doing (possibly all) of the following:

1. **tracking** the movement of the puck
2. **watching** for signs of it slowing down or changing direction.
3. **hoping** it will stop near the target
4. **willing** it to stop near the target
5. **being aware** of the imminence of the result.

Our second question will be:

- What elements in this situation are essential to feeling suspense?

[1]We expressly neglect the presence of the ice brushers clearing the ice in front of the puck as it slides forward.

We will group our answers to this question under two main headings: tracking and timing.

3.1.1 Tracking

Tracking the puck is roughly equivalent to estimating the next position of the puck and verifying that it goes there. As we watch the sliding puck, we are *constructing* its path through the curling space. Some kind of tracking is essential for suspense to exist. This *path construction process* has certain characteristics:

Entity and event identification

The object being tracked must have certain unchangeable features with which it can be identified. The puck has its hardness, solidity and shape. Characters in a story may have their physical appearance, their name, their family, their age and so on.

In a similar way, we need some way of identifying and understanding specific events that happen to the objects or characters in the given situation. Different things happen to the puck at different moments along its path, just as they do to a character in a story.

Describing and summarising the future

There is a way to create descriptions which summarise possible future behaviours. For the puck, we can use phrases such as ‘it’s going too far to the left’, or ‘it’s heading for that gap’.

For a character in a story, one seemingly efficient (and perhaps inevitable) way of summarising a character’s future behaviour, aside from their current physical actions, is to imagine a series of *goals* or *plans* that the character is pursuing.

Watching for... the role of prediction

As the puck slides forward, we are looking for signs that will confirm or disconfirm our current descriptions of the future behaviour of the puck. These could be for example: ‘it’s slowing a lot now’, ‘there’s some smoother ice coming up which will slow the puck less’ and so on. ‘Watching for’ is roughly equivalent to having a *recalculation program* ready to be applied to update the estimated stopping point of the puck. The program is run as soon as we observe a discrepancy between the predicted and observed values of the puck’s path.

Similarly, for characters in a story, we are expecting, are attuned to, are *watching for* the appearance of certain signs which will enable us to make more accurate updated models of what will happen to them. Often we are looking out for signs that will help solve the character’s problems, or events that could affect the character’s goals.

Updated descriptions of a possible future are being generated continuously in suspenseful situations. There is also a constant choosing of the most valid description from the available possibilities. And of course, as we watch for what might happen, we are always also trying to estimate *when* it will happen.

Consistency of inferences

There must be some process which maintains the logical or causal consistency of the past and predicted parts of the path in question. For the sliding puck, this could be the association of say, early spinning with later curving, or early speed with later distance.

For characters, this could involve physical, biological and psychological consistency between different life phases, say, someone who is good with their hands becomes a gardener.

This suspense-orientated description of a sliding puck leads to some other consequences, which we will now examine.

3.1.2 Timing

Feeling suspense is something that happens in real time. The constraints on attention and working memory play a role in how it is felt. In this respect, recent research on short video-clips has shown that the brain integrates incoming sensory information as it unfolds over time periods of 2 to 3 seconds (see Fairhall et al., 2014). We now examine certain aspects of suspense which have to do with timing.

'Neither too quick': the limits of recalculation

The whole tracking process requires time to be carried out. Changes and updates take time. This means that the events being tracked must not happen in too quick a succession if we want suspense to be maintained. If events occur too quickly, then we will not have enough time to maintain the consistency of our understanding of the new situations and we will therefore not be able to generate new sets of predictions.

In curling, one of the attractions of the sport could be that the puck does not travel *too quickly*. Spectators have enough time to track and recalculate the puck's path from even small deviations as it slides across the ice. We can compare this situation with that of a footballer kicking the ball into the goal. Unless the shot comes from quite far out, we usually don't have time to update the suspense we feel *as we watch the changing course of the ball* *along its path*. We simply feel suspense about whether the ball goes in the net or not.

For characters in a story, similarly, sequences of events where too many things are happening at the same time may produce feelings of bewilderment, and possibly bedazzlement, but probably not suspense about individual events.

There would appear to be an upper bound on the speed at which events can occur for suspense to be maintained. This may be one of the reasons that some film-directors use slow-motion techniques at culminating moments in narratives where many things are happening at the same time.

'Nor too slow': no need for recalculation

In a similar way, even if they have not changed, predictions about future events must regularly be *remembered* or *re-evoked* for their effect on suspense to be maintained. If they are not re-evoked from time to time, they go to the back of the viewer's mind and their suspenseful effect wanes. Apart from telling and thus confirming predicted events, one way to re-evoke a prediction is, of course, to put it to the test by creating counter-predictions which might invalidate it. Such counter-predictions force the original prediction back into working memory, potentially re-triggering its suspense-producing effects.

To summarise, we can say that if counter-predictions or confirmations of predictions are not regularly produced from time to time in a suspenseful situation, then the suspense drops because we become either uninterested or certain of the result.

In the curling case, such a situation might occur if the puck's travelling speed were, say, a tenth of what it is. Spectators would recalculate the puck's path much less often, and during the time between each recalculation, they would perhaps be free to think about other things.

In a story situation, a similar situation could occur if a character's goals were left unmentioned or unchallenged for long sections of the story. Any suspense generated by these goals might then fade into the background.

Many stories regularly imperil important predictions by creating conflicting predictions which could potentially annul them. A secret agent on a rescue mission might thus regularly undergo set-backs which have the potential to jeopardise the whole mission.

Suspense and imminence

In the light of the preceding discussion of time-scales, we now examine the concept of imminence, the *chronological aspect* of suspense.

Imminence is realising that the all or nothing moment is approaching. We concentrate on the moment as there is no going back, there is a realisation of the one-way nature of what is happening. Imminence is a signalling that one must prepare for a cognitive and/or emotional upheaval in one's understanding of a situation. It is a function of *when* events are predicted to occur:

- *When will the moment arrive, for which I must be ready (cognitively and emotionally), when one whole set of possibilities will no longer be valid and a new set will appear?*

As we discussed briefly in 1.1.3, it appears that bodily reactions increase by degrees according to the reduction in the perceived distance from the source of a threat. This reduction in the perceived distance can of course be translated into an increase in the *imminence* of the threat: the *closer* the danger, the *sooner* it is likely to reach me. We can therefore extend the concept of perceived physical distance from a danger to one of perceived distance *in time*[2].

The time-scales of suspense

Different time scales can allow us to feel suspense about different things at the same time. We feel suspense increasing as the upcoming event becomes more and more imminent. For the puck, we may feel both long-range suspense about the final result of the puck slide, and short-range suspense as we observe small changes in its trajectory and notice whether it is attaining intermediate goals.

We think that there could also be a kind of recursive 'suspense about the suspense', or suspense about *the suspense we expect to feel* during the very last phase of the puck's trajectory. We might thus feel 'low-imminence suspense' or 'long-range suspense' about a certain period to come within which we expect the suspense and excitement level to be much higher. Once we enter this final highly suspenseful phase, again we can have a kind of suspense about a yet higher suspenseful phase as the actual moment of the final outcome approaches. Thus there may be a kind of *recursivity* of suspense levels: suspense is a type of emotion about an emotion (about an emotion)...

As another example of this, we can think of a football match where we may feel long-range suspense about the final result before and during the match, and short-range suspense about particular attacking movements on the pitch during play. We may also additionally feel suspense about the suspense we will feel in the final minutes of the game.

[2] We speculate that there maybe also be a kind of *compression* of time-consciousness equivalent to the brain functioning more quickly. This could also be the inspiration of the slow-motion techniques used in film.

To summarise, in typical suspenseful situations, we can distinguish the following three types of suspense:

- Long-range, low imminence suspense: 'which team will win the match?'
- Short-range, high imminence suspense : 'what will happen in the current attacking movement?'
- Final high-imminence suspense: the moment where the preceding two suspense types coalesce.

3.1.3 Summary of useful features

From this study of curling, here is a summary of some features which we would appear to need to create a model of suspense in narrative:

- Coherent identification of specific objects or characters in the narrative
- Summaries of future events
- Watching for specific future events
- Consistency of inferences about events
- Sequences of events that change neither too quickly nor too slowly
- A measure of the degree of imminence (of resolution) of the suspense felt
- The possibility of different time scales of suspense being active at the same time.

We now examine some models of narrative comprehension to show how these features can be included in a model of suspense grounded in the narrative comprehension paradigm.

3.2 The narrative thread: a data structure for suspense modelling

The goal of this research is to find an appropriate and simple way to model narrative which allows effective modelling of suspense. We now integrate elements from the psychological literature on narrative comprehension, Brewer and Lichtenstein's approach to stories and some theoretical narrative constructs of our own to derive **a model of narrative comprehension which allows for the modelling of suspense**.

3.2.1 Discretisation: splitting stories into 'simple' events

In our analysis of the sliding puck, any of our attempts to describe in words the puck's *continuous* sliding movement will inevitably pick out *discrete* parts of its path: 'it's just gone past the 10 metre mark', or will be *predictions* of the puck's future movements: 'it's slowing down'. To tell a chronological story in words, we need to make choices about how big such chunks should be. We could choose for example, each individual sentence in the following:

> The puck slid forward another 10 cm. It started to spin slowly. It passed the 5 m mark. The rough ice slowed it down. It stopped spinning. It slid past the 10 m mark. It slowed down once more. It drifted to the left. It stopped 30 cm away from the target.

or perhaps just this terse summary:

> The puck slid 15 m and stopped near the target.

To inform our *design* decisions about the degree of detail or the span of the events we use in our model, we will refer to our previous review of event segmentation in chapter 2 (see 2.3.2)and use Zwaan's protocol as our basis (Zwaan et al., 1995). This protocol uses the following criteria for event segmentation:

- new space
- new interaction
- new subject
- new cause
- new goal

Examining the protocol, we see that the former more detailed description is more appropriate. Within the framework of suspense modelling, the crucial question is also a *causal* one: 'did something just happen that may have an influence on the final outcome?'. This protocol provides a way to answer the condition of summarising the future that we evoked previously.

To summarise, to extrapolate events from a storyworld, we can use the event segmentation protocol on any stories which occur in this storyworld. Translating this protocol into a series of constraints, we obtain that an event should have a unique temporal reference, space, type of interaction with a given object, subject and activity. Any event that appears to have more than one of these elements should be split into several events.

3.2.2 Fabulas and Storyworlds

Using our curling situation, we can clarify the relationship between *fabula*[3] and *storyworld*. A storyworld is the imaginary space in which a number of similar stories can be told. It has the following characteristics:

[3]For pratical and expressive reasons, from now on we will treat the words fabula, syuzhet as standard English words which do not need to be italicised or capitalised.

- It can be described by a set of constraints and rules regulating possible sequences of events implicating the objects and characters which populate it.

- It can be used to generate a set of fabulas.

Under this definition, a fabula is a chronological list of possible events from a storyworld. Once chosen, fabulas can be *told* in different ways: by revealing or hiding some of their events, or changing their order of presentation. This is the *syuzhet* of the Russian formalists.

Let us assume that the curling events we have described are not taking place in the real world but rather in a **Curling storyworld**. Further we will assume that there are only two possible fabulas in this storyworld that can be told: fabula A, where the puck reaches its target, and fabula B, where the puck misses. As we observe the unfolding events, we want to know which of the possible fabulas is being told: A or B. Let us suppose that the first n events in each of these two fabulas, however, are exactly the same (at least for a given degree of perceptive accuracy). While the first n events are being told, we are constantly tracking the puck and projecting into the future two possible fabulas which seem coherent with the movements of the puck thus far.

Now certain events in these fabulas are incompatible with each other; the puck cannot stop both near and far away from the target, and these conflicting events mean that there is suspense. Sooner or later an event will occur which will differentiate the two projected fabulas, thus disallowing one of them. When the $(n+1)^{th}$ event is told, we immediately know which of the fabulas is actually being told and will therefore succeed. The future of the puck is decided and suspense dissolves.

In this reading of the story comprehension process, we feel suspense while we are waiting for an event which can *disambiguate* which of the possible fabulas is actually being told.

3.2.3 Data structures for causal consequence inferences

Once we have identified the objects or characters and the possible events that can occur in a given context, then we must consider how we model the description and summarisation of future events. Inferring possible future events, in other words, forming predictions, requires knowledge. Such knowledge can be explicit or implicit and take the form of concepts, categories, schemata, rules, simulations and scripts.

Referring to our review in the previous chapter, our starting point for inferential process modelling is the *constructionist model* of narrative comprehension (see 2.3.2). The constructionist theory claims that the only inferences that are made during narrative comprehension are those needed to construct a coherent explanation of the narrative content. It further stipulates that some classes of inference are not constructed during comprehension because they require too much time or cognitive effort. These include logic-based inferences, detailed elaborations and distant causal consequences. We relax the constraints of this model a little and include in our model of suspense parts of the *prediction-substantiation model.* This step enables us to integrate the general use of *causal consequences* as a type of inference that can be generated during narrative comprehension. Two conditions, both emphasising the *ease* with which knowledge structures can be retrieved from long-term memory, constrain the possibility of a causal consequences inference:

- a strongly supporting context

- a strongly directive context

We identify the following three classes of inference which have these qualities:

Simulations: these are mechanisms for projecting certain states of affairs in the story world into the future, perhaps until they run up against some physical limit. Some examples could be applications of naive physics: 'the water is rising quickly and will reach the ceiling in a few minutes'.

Rules: these are general formulae or very basic deductive reasoning mechanisms which can be applied one-off to given situations to make predictions: 'if you drop an object, it falls down'.

Scripts: these are easily accessible knowledge structures in the form of sequences of different events which usually occur in a given order. They can vary considerably in complexity.

This research will mainly concentrate on the use of *scripts* for suspense modelling in narrative. We leave the integration of simulations and rules for future work. We now examine briefly some different types of script, ranging from the simple to the complex.

Script complexity

Perhaps the simplest type of script we can call a *narrative thread.* Narrative threads are made up of a fixed sequence of events which typically occur one after the other in a given storyworld. Narrative threads are completely linear and unambiguous and have one unique outcome and thus *one unique value* in a storyworld. This value is equivalent to the evaluation by the reader of the state of the storyworld *after the last event in the narrative thread has occurred.*

We can symbolise a narrative thread in the following way:

$$A \rightarrow B \rightarrow C \rightarrow \ldots \rightarrow Z \tag{3.1}$$

Plot fragments are a more complicated than narrative threads, and contain several choice points where different sub-threads can be followed[4]. They nevertheless still only have one possible end result, so their value in a storyworld is unambiguous. The choice points of a plot fragment are all directly connected to one basic chain of events. They can be symbolised as follows:

$$A \rightarrow ((B \rightarrow C \rightarrow D) \text{ or } (E \rightarrow F)) \rightarrow G \rightarrow \ldots \rightarrow Z \tag{3.2}$$

Situation scripts differ from plot fragments in that they contain several possible end-points with different outcomes in the storyworld. They can have multiple branching points, but they may also have some events which contain other embedded scripts or plot fragments. They usually describe events which typically often occur during the same time period or at the same place and pack into a single object a group of events which are *associated* with each other in memory[5].

We summarise the differences between these three knowledge structures as follows:

- Narrative threads: no branching, single outcome
- Plot fragments: some branching, single outcome
- Situation scripts: wide-ranging branching, multiple outcomes

In this research, we will restrict ourselves to *narrative threads* which have absolutely no branching. This design choice has the aim of reducing

[4]One well-known example is the model proposed by Lebowitz (1985).
[5]One well-known example is the model proposed by Schank and Abelson (1977).

complexity in line with our main research goal which is to put forward a *possible* model of suspense.

We now link up the preceding considerations with Brewer and Lichtenstein's model of story.

3.2.4 Integrating Brewer and Lichtenstein's concept

Brewer and Lichtenstein's formulation of suspense (Brewer and Lichtenstein, 1982) requires the presence of an Initiating Event (*IE*) which predicts the possibility of a particular Outcome Event (*OE*). To create suspense other events are placed between *IE* and *OE*, and *IE* and *OE* are ordered chronologically. In Brewer and Lichtenstein's terms, the appearance of the *IE* in the narrative triggers the existence of the *IE*-*OE* pair in the mind of the reader.

If we try to make the concept of an Initiating Event and an Outcome Event more precise, two questions arise:

- How does this *IE*-*OE* pair create suspense?
- What exactly links the *IE* to the *OE*?

The answer that our model provides to the first question is that suspense is based on the idea of:

- detecting two or more *conflicting* predictions.

Brewer and Lichtenstein's Outcome Event must actually correspond to a *set* of (at least two) *predicted but conflicting* events.

The simplest possible answer to the second question would be:

- a fixed linear sequence of events leading from the Initiating Event to the Outcome Event

But where would the information come from to create this sequence of events? And how could we formally define such a structure?

3.2.5 Narrative threads characteristics

We now combine the previously discussed requirements on knowledge structures for narrative comprehension with Brewer and Lichtenstein's story model. These requirements taken together can provide the basis for a definition of the *useful* size and complexity of a script:

1. It must have *one* starting-point and *one* end-point. This is the concept of an Initiating Event and an Outcome Event from Brewer and Lichtenstein.

2. It must have a *clear path to completion.* This is the condition on causal consequences from constructionist theory.

3. It must be *interruptible.* Suspense requires the potential for a conflict between events which produces a degree of uncertainty.

We will use the knowledge structure **Narrative Thread** to create such a linear sequence of events. Thus, instead of an *IE-OE* pair, we use the following:

$$IE \rightarrow Event_1 \rightarrow Event_2 \rightarrow Event_3 \rightarrow ... \rightarrow OE \qquad (3.3)$$

In fact, we put forward a **narrative thread** as the appropriate structure for suspense modelling as defined under the terms above. The narrative thread has the necessary clear final result or outcome in the storyworld that we need for suspense and, according to the constructionist model, is also habitually and easily generated during narrative comprehension.

Narrative threads can be thought of as ordered simple lists of events which are likely to follow each other in a given storyworld. The events in a given thread are ordered according to their typical sequential occurrence in the *narrative time* of the storyworld. Of course, causality often plays a role, but in general, threads can be informed or governed by a variety of inferential mechanisms: scripts, models of story characters involving beliefs, goals and desires, principles of naive physics, narrative traditions, and so on.

Our claim is that whatever techniques we use to model the storyworld, we can translate the available information and inferences into the narrative thread form. Narrative threads can be thought of as lying somewhere between the storyworld and the fabula: they give information about the storyworld but also indicate likely sequences of events in a given storyworld. Furthermore, they can be used to construct parts of stories.

We can give some characteristics of narrative threads which concern their use during the reading of a story.

Online computation

Our claim is *not* that all narrative threads are a standard part of readers' knowledge when they start reading a story. Rather, we claim that most such structures are *assembled* by the reader from a variety of causal and intentional information sources. The activation and construction of narrative threads to model a storyworld is carried out during narrative comprehension.

Consistency

Following the distinction made in the GLAIVE project (Ware and Young, 2014), we can distinguish a narrative thread's *causal* and *intentional* links. GLAIVE introduces conditions on causal chains and intentional paths which are mostly concerned with maintaining consistency. These conditions can be applied to narrative threads and we summarise them as follows:

- No event in a causal chain can negate the preconditions of another event in that chain.
- A character must consent to all steps in a intentional path and intends the final effect of the last step during all the preceding steps.

Stability

A narrative thread is then a sequence of events which has a type of *internal consistency* over all its events. Major updates and changes in the content of specific narrative threads can of course also occur while reading. However, the claim we are making is that narrative threads have *sufficient stability* during narrative comprehension to have a high utility in modelling suspense. This feature of narrative threads is also what distinguishes them from simple pairs of connected events.

Content is separate from function

Further, we claim that it is not necessary to know exactly *how* a range of information sources were used to construct some narrative threads for them to be used successfully to model and predict suspense. We merely need to know that such narrative threads *can* be constructed. An essential characteristic of our model is to postulate a separation of the inferential or causal *information sources* of narrative threads from the *formal structure* of the actual ongoing suspense process that they generate and maintain.

3.2.6 Consistency of inferences: the need for interruption

In order to model suspense, we need an additional feature: there must be some way for events to interrupt or conflict with the events in a given thread. Narrative threads must be interruptible. We can express this relation in the following way:

$$\text{event } A \overset{disallows}{\Longrightarrow} \text{ event} B \tag{3.4}$$

This is to be interpreted in the following way: ‘if Event A occurs in a story, then Event B can no longer occur in this story’.

Alternatively, if Event A disallows Event B and Event B is a part of Narrative thread N, then we can say:

$$\text{event } A \overset{interrupts}{\longrightarrow} \text{ narrative thread } N \tag{3.5}$$

This means: ‘if Event A occurs in a story, then Narrative thread N can no longer be **active** in this story.’

For a given storyworld, we can thus construct a set of event-pairs where event A **disallows** event B, which we notate thus: $(\text{event } A, \text{event } B)$. The derivation of these event-pairs depends of course on information about the storyworld. As before in the case of the narrative thread content, the set $\mathbb{D}$ of disallowing event pairs can be informed by a wide variety of sources.

3.3 Competing narrative threads

The events in given narrative threads may or may not be known by characters in the story. Only the reader is party to all the different threads and can predict their interactions. Indeed the reader must know about the possible conflict points between different threads in order to feel suspense. However, different threads will have a different degree of importance for the reader, and the reader's attention to a given thread will also fluctuate. To adequately combine the effect of different threads and to take into account the fluctuations in a reader's attention as a given story is told, we need to build the following elements into our model:

- A measure of the relative *importance* of different narrative threads, and
- A measure of the varying degree of *attention* that each narrative thread receives.

3.3.1 Importance

Whilst a reader is reading a given story, there may be several suspenseful situations present at any one time. Relatively unimportant suspenseful situations may coexist with life-or-death situations. The importance of a suspenseful situation for a given reader will also often depend on their emotional involvement with any of the characters. They may have a low or high, positive or negative emotional involvement. A negative emotional involvement could be ascribed to the 'baddies', and positive involvement to the 'goodies'. We need a way to model the relative degrees of importance of each thread, if we want to obtain a suspensefulness measure.

Our goal is to define suspense in a way that does not depend on specific world knowledge, but how can we nevertheless account for such differences in importance? As our main research focus is on identifying some minimal requirements for modelling suspense, in our model, we will *presuppose* the existence of a mechanism which can ascribe a relative emotional importance value to any event, possible or actual, in a given story. In this way, we can ascribe an importance value to every narrative thread, because, as we discussed in 3.2.3, the value of a narrative thread is equivalent to the value the state of the storyworld after the *last* event in the narrative thread has occurred

The importance of an event under this definition will depend on what has already been told in the story. At this stage in our research, however, we will be making the assumption that the importance of all events (and therefore the importance of all threads,) stays the same wherever we are in the story. This use of one importance value to encompass a multitude of factors enables us to simulate the modelling of emotions and at the same time keep a clear-cut separation between the suspense algorithm and domain- or story-specific information. In this way, we can concentrate our research on the structure of the information flow and its relation to suspense.

As we have already discussed in 2.5, suspense *can* exist without the explicit presence of human characters. Ascribing an importance value to an event is therefore a very general procedure which may depend on a wide variety of factors. For instance, how would we determine the importance value of a ball falling off a table? It seems that merely provoking any *irreversible* change in some feature of a storyworld carries with it a degree of importance. As far as many stories are concerned, however, the importance of events is measured in relation to the viewpoint of some human character. We will therefore mostly base our ascription of importance values on the following two factors:

- The current level of sympathy (or antipathy) towards a character involved in an event,
- The perceived desirability (or undesirability) of the event in relation

to that character.

The simplest possible modelling of importance which takes into account the two factors above is use the product of two numbers. We choose the following scales:

- a sympathy level ranging from *–1* (hate) to *0* (neutral) *+1* (love)
- a desirability level of an event for a character ranging from *–10* to *+10*

An event which is deemed to be 'bad for a bad character' would therefore get the value $-1 \times -10 = +10$.

3.3.2 Foregroundedness

In addition to a measure of the relative importance of the narrative threads, we need a measure for the *degree of foregrounding*, or the *foregroundedness* of a thread throughout the story-telling process. If a given narrative thread is not mentioned or in some way evoked for a certain time during the telling of a story, then we assume its effect on suspense will drop, because it will no longer be so present in the reader's mind. Moreover, at the same time, other narrative threads will of course be active and competing for the reader's attention.

Of course, as soon as a new event in a narrative thread is mentioned, the narrative thread in question comes again to the foreground, regaining all its potential for suspense-creation.

The term 'foregroundedness' is also used in discourse analysis as part of the foreground-background dichotomy (see for example Virtanen, 2004, p. 100–101 for a review of the phenomena to which the term is applied). Our use of the term, however, defines the degree to which a narrative thread is in the forefront of the reader's mind, and is roughly equivalent to *recency*. Recency has been extensively researched in the psychological field as an important factor influencing memory (see for example Jones et al., 2006). In our model, foregroundedness is linked to recency of mention, in that the level of foregroundedness of a thread depends on how recently the particular narrative thread was evoked or mentioned in the story. If a narrative thread has just been mentioned directly in a story, then it will have maximum recency and foregroundedness. However, in our model, a thread may also be mentioned *indirectly*, when other threads predict some of its upcoming events for example, and this too will bring it to forefront of the reader's mind and increase its foregroundedness.

We will mostly use foregroundedness to model the forgetting of a narrative thread that occurs if it is not referred to for a while during the telling of the story. If a narrative thread is activated just once and never evoked again, a model of the variation in foregroundedness might have an initial high peak value before going down and reaching a stable plateau level. When it is re-evoked, then its foregroundedness value goes again back to a maximum level.

There may be a threshold effect such that every new evocation of a thread weakens the forgetting process so that certain threads can no longer be forgotten even if they are very rarely mentioned. Be that as it may, we will use a simple one-dimensional mathematical model of the degree to which unmentioned events leave working memory, and therefore the degree to which they no longer influence suspense.

3.3.3 Imminence

In addition to these measures, we include an additional factor in our model: imminence[6].

Every narrative thread has the potential to create a **change in the storyworld equivalent in importance to its importance value**. It achieves this notably by 'succeeding', which happens when the thread's last event occurs in the story. Depending on how many untold events remain in the thread before it can be completed, we say that the completion of the thread is more or less *imminent*.

As soon as we have multiple narrative threads present in the modelling of a story, however, there is however, another kind of imminence at work. For a given narrative thread Z, there may be an event υ in another thread Y which could soon occur in the story and which would *disallow* some event δ in Z. In so doing, the occurrence of event υ would disallow the whole thread Z and the change in the storyworld that it could bring about would no longer be attainable. There is therefore the imminence of *interruption* of a thread as well as the imminence of its *completion*. Accordingly, each narrative thread will produce suspense due its potential *completion* but also due to its potential *interruption* by disallowing events in other threads. We will use the following terms:

- Completion imminence
- Interruption imminence

We can informally summarise these two kinds of suspense by asking the following questions:

[6] As we discussed briefly in 1.1.3, it may be that the imminence of a suspenseful situation increases the amount of physiological, emotional and not just cognitive reaction that it provokes.

- how soon might a narrative thread succeed?
- how soon might a narrative thread fail?

A given narrative thread could for example be a long way from completion but close to failing due to imminent interruption by a *different* narrative thread, or vice versa. We will assume that these two types of imminence are *independent*.

Once a thread has 'succeeded' then it can no longer be disallowed. Completion imminence is our attempt to model the *dispelling* of the possibility of a given thread being disallowed by other hitherto unknown or hidden threads in the storyworld. Completion imminence is based on an assumption that the storyworld model is always incomplete.

So, in our model, the suspense that a given narrative thread contributes to a story also depends on how many narrative threads could disallow it. Or, informally, a narrative thread creates suspense in a story by 'threatening to disallow' other narrative threads.

3.3.4 Modelling imminence

Modelling the relationship between consecutive events

One could model the relationship between the different events in a thread using transition probabilities. Schematically this could be shown in the following way:

$$A_1 \xrightarrow{p_1} A_2 \xrightarrow{p_2} \dots \tag{3.6}$$

where p_n is the probability that event A_{n+1} will occur if event A_n has occurred in the storyworld.

Of course, such probabilities may be independently obtainable from some real-world source. However, the projecting of real-world probabilities on narrative situations is a complicated process. A sequence of events that is very rare in the real world may well be very common in stories. As an example of this, consider a man walking into a bank. In everyday life, this is a banal occurrence. In the context of a story however, the simple fact of telling this event might lead a reader to expect a bank hold-up story.

Our contention is that real-world information, although serving as a useful basis for extrapolating likely sequences of events, may not be appropriate for precisely modelling probabilities of events occurring in *stories*. We simply do expect surprising and highly unlikely things to happen in a story context.

These considerations have led us to focus on using a much simpler linking relationship between the events of a thread. Our choice has the added advantage of reducing the real-world knowledge we need to construct our model.

We make the following assumption: if a given narrative thread does actually correspond to the fabula being told, then the predicted upcoming events in this narrative thread are expected to happen (eventually) with a probability equal to 1. The only possibility for an upcoming event *not* to occur, therefore, is catered for by the possibility of the thread being disallowed by some other event and this eventuality would mean of course that the narrative thread in question does *not* correspond to the fabula being told.

This constant probability of transition between all consecutive events in all narrative threads that we have chosen, could take on other values, less than *1*. While such a step might be useful in modelling the real world, we are uncertain as to its value in modelling narratives.

Imminence and narrative time

Assuming a constant probability between transitions is one design simplification we use in our model. Another related feature is the assumption that the *perceived narrative time* which needs to elapse between two consecutive events in a thread is approximately constant.

Perceived narrative time is a measure of narrative time from the reader's point of view. It is an idealisation of how soon the reader expects a given event could occur in the ongoing narrative process. It is not the same as the time in the storyworld.

The assumption that the perceived narrative time between events in a thread is constant leads to the possibility of a simple measure for imminence. In our narrative model, we use the following general measure of the *imminence* of an event:

- the *imminence* of an event is proportional to the number of events that can still occur in a narrative thread containing the event before the event itself occurs.

Imminence and Probability

Our simple narrative thread structure thus couples two functions:

- Narrative Imminence

- Probability of occurrence in a narrative

A *decrease in the probability* of an event occurring in the storyworld can be simulated in this model by *increasing the number of events* before it in the thread in which it occurs. However, in our model, this procedure also *lowers the imminence of its occurrence.*

As we are using a fixed probability for all transitions between events, the only way to reduce the probability of a given event occurring is to increase the number of events between it and the current event. This suggests that the event segmentation maybe need to be modified; conceivably, a narrative thread might require additional events to be inserted in order to more closely model readers' reactions.

We have made a design choice to model imminence in this particular way. There are of course other possibilities for the modelling of perceived narrative time. They lie however outside the scope of this research and are left to future work.

3.4 Suspense and Curiosity

Let us imagine the situation where someone suddenly walks into a room with a huge smile on their face. We can imagine the questions and thoughts that the observers of this event might have; perhaps a first question based on past events - 'what happened to this person to make them so happy?' - could lead to the thought 'in any case, it must be important' and also to questions about potential future events: 'what will this person do now?', 'will it be exciting or dangerous?' and so on. Many plausible and different *answers* to these questions could be entertained: 'they've just won a lot of money', 'they've fallen in love', 'they've had too much to drink', 'they're happy about their imminent plan to do a surprise song-and-dance number for the assembly'... In our approach, these potential explanations are modelled as a set of narrative threads, all of which contain the 'smiling event' in their sequence of events.

In some pilot experiments we conducted, we found that the suspense evaluations given by participants for situations similar to this one increased in leaps which were much greater than those predicted only by imminence, that is the increasing proximity to the completion of a thread. We hypothesised that one possible explanation for this phenomenon could be that new narrative threads might need to be fully ‘confirmed’ before they can have their full effect on suspense. If a narrative thread Z started out with a *low* level of confirmation due to the competition from other conflicting narrative threads, then, the elimination of these other threads could create big leaps in the confidence we have in Z.

Going back to our example, as events in the room start to unfold - ‘hey, he’s got a funny hat’ - and all but one of the tentative explanations for the smile are eliminated - ‘so it *was* the song-and-dance number’ - the confidence in the remaining explanation will go up sharply. We may then arrive at a classic case of conflict-based suspense: ‘I hope he’s not going to use me as a volunteer’.

One simple way to model this phenomenon is to suppose that at first, all threads are activated with a low confidence level, and that every time one of the events in the thread is told in the story, the reader’s confidence that the thread is applicable to the story increases. To simulate this effect, we first experimented with the following *confidence function*:

$$\text{Confidence} = 1 - \frac{1}{(2 \times \text{ Evocations })} \tag{3.7}$$

where *Evocations* is the cumulative number of times the narrative thread has been evoked due to one of its events occurring in the story. For the first four evocations of a thread, this function produces the following factors: $0.5, 0.75, 0.83, 0.88, \ldots$. After 5 or 6 evocations, the function has little further effect on the suspense level as it stabilises close to 1. At this stage, we might consider that the thread has been 'sufficiently confirmed'.

This simulation of thread confidence seemed however rather *ad hoc*, and this led us to develop a different theoretical derivation for this phenomenon more in line with the rest of our theory. The concept we developed draws some of its inspiration from Brewer and Lichtenstein's work, and is based on the idea of curiosity, another of the key features in their modelling of stories. We call it *Revelatory suspense.*

3.4.1 Two different suspense mechanisms

We claim that two fundamental mechanisms occur in stories to create suspense. We call these *conflict-based* and *revelatory*:

- **Conflict-based suspense** occurs when two narrative threads appear to lead to two incompatible events; only one of these events can occur in the given storyworld. Furthermore, a big difference in story outcomes is expected depending on which of these two events does actually occur. An example would be a race between two runners.

- **Revelatory suspense** occurs when many different narrative threads could have led to a particular event being told in the story. Our situation of someone walking into a room with a huge smile on their face exemplifies this type of suspense. We are uncertain as to which narrative thread is the right interpretation of our event. These narrative threads also mutually conflict with each other[7].

To summarise, incompatibilities or conflicts between events predicted to occur *after* the current event ('in the future') produce *conflict-based*

[7]In a sense, even conflict-based suspense is a type of revelatory suspense. Which of the two incompatible events actually occurs in the storyworld is known by the author: it is 'in the past'. But this is also information that we know that we do not know; that is why we read the story.

suspense whereas incompatibilities or conflicts between events presumed to have occurred *before* the current event ('in the past') produce *revelatory suspense*. Recast in the terms of our narrative thread model, *conflict-based suspense* exists between two confirmed threads, and *revelatory suspense* exists as the reader disambiguates between partially unconfirmed and incompatible narrative threads.

3.4.2 Revelatory suspense

Curiosity in Brewer and Lichtenstein's model of narrative is about the past; we wonder why something happened, or we know that something happened in the past about which we do not know enough. The notion of *revelatory suspense* is a combination of suspense and curiosity. It is *suspense linked to an event about which we are curious.*

For some events in a story (or in real life), we can have the intuition that there is something that we do not know about the event, that something is hidden from us. This intuition can be triggered in many ways. Perhaps an unnecessary detail is lingered on; we wonder why. Perhaps a character behaves in a strange way; we wonder what it is that we do not know about him. In other words: *we know enough to know that we don't know enough.*

As we go through the story, we acquire more information about the story and the storyworld and eventually come across information which fills in the gaps in our knowledge, and makes one specific interpretation of the hitherto strange behaviour more and more plausible to us. Eventually, we have a clear interpretation of the strange event, and thus of what might happen in the future. This epistemological gap-filling process is suspenseful in itself: we know that each moment where we find out new information is likely to be an important one.

In this way, we see that a given narrative thread can have varying degrees of confirmation or, as we will label this, *confidence*. However, computationally, such 'known unknowns' are hard to model directly. We show the mathematical approach we adopted in 4.3.

3.5 Suspense and Surprise

Our narrative thread model also provides a way to model *surprise*. If thread A has many told events and thread B has far fewer or even no told events, then the reader will feel surprise if an event in the story occurs which suddenly disallows thread A and boosts confidence in thread B. We say, the reader was expecting events from thread A to occur, but events from thread B occurred instead. In an extreme case of surprise, the event in thread B which disallows thread A is the very first event of thread B to occur in the story.

Note that if threads A and B were equally confirmed threads - the situation of two runners in a race for example - then, although we might feel suspense, we will not feel surprise when one of them finally succeeds. For surprise to exist, there must be at least one (partially or wholly) unconfirmed thread which becomes confirmed.

The precise modelling of surprise is however outside the scope of this research and is left for future work (see 7.2.6 for further discussion of surprise and its relation to the narrative cycle).

3.6 Data sources for narrative inference

3.6.1 Event chain modelling and derivation

Part of our approach to suspense is to disassociate the content of the narrative threads from the existence of suspense. We can formulate this in the following way:

- if narrative threads can be built from different sources of information and
- two or more events in these different narrative threads are in conflict,
- then we have a case of suspense, independently of the sources which inform the threads in question.

But how can we obtain the necessary information to build our narrative threads? The design choices we have made for our model attempt to take into account current automatised or semi-automatised knowledge acquisition techniques.

Chambers and Jurafsky (2009) presents an automatic harvesting technique from corpora which enables the acquisition of information about typical linear sequences of events called *event chains*. These are based on verb occurrences and ordering in real-world texts. Their work thus enables a derivation of **temporal precedence constraints**. In our terminology, such an event chain is equivalent to a narrative thread.

Li et al. (2013) has taken this approach a step further, and describes a system where plot graphs describing or resuming story situations can be derived using crowd-sourcing techniques. In the example they give, a series of events that could typically occur in a bank robbery are linked together to create a general plot graph of this situation. Two different types of link between two events are used: *precedence relations* and *exclusion relations*. The distinguishing feature of this work is that it can generate information about **mutually exclusive events** and this type of information is also an essential part of our suspense model.

3.6.2 Conditions on narrative threads

As we have discussed in 3.2.3, aside from these automatically crowd-sourced scripts, we can also create narrative threads from simulations or rules. We may also want to create story-specific narrative threads which integrate information from very different character models (based on goals, plans or emotions for example).

In all cases, for all types of narrative thread including simulation- and rule-based threads, we can distinguish *causal* or *intentional* links, and we can apply consistency definitions similar to those used by GLAIVE (Ware and Young, 2014), which we have described in 3.2.5. Such conditions could ensure that information from a wide variety of sources can always be transformed into the narrative thread form.

3.7 Summary

In this way, by combining an analysis of a simple suspenseful situation such as curling with aspects of constructionist narrative comprehension theory and Brewer and Lichtenstein's model of suspense, we have built up an informal description of a model of suspense in narrative based on the following points:

- Highly constrained sequences of events called narrative threads which are predicted to occur in a given storyworld, can be used to model suspense in a narrative. Narrative threads are simple linear structures which can be generated by automated processes.

- Storyworld knowledge about events that conflict with each other in the form of a set of disallowing event-pairs is necessary to allow for the possibility of narrative threads interrupting each other.

We can derive a value for the suspense that a given narrative thread creates at any point in the telling of a story by combining the following four variables:

- **Importance**: 'what is at stake' in the storyworld in the case of the completion of the narrative thread,
- **Foregroundedness**: a measure of how present the narrative thread is in the reader's mind
- **Imminence**: the relative event-based proximity for a given narrative thread to be completed, (completion imminence) or interrupted (interruption imminence)
- **Confidence**: This is not the confidence in a particular story outcome, but rather the confidence that a particular narrative thread provides the *right* interpretation of some events in the story.

We claim that increases in any of these variables will increase the suspense reaction.

Chapter 4

A mathematical model of suspense

The mathematical formulation of our suspense model has the following parts:

- The definition of a storyworld
- The definitions of a story and a fabula and their interaction with a storyworld
- A description of how a story is told
- The definition of the reader's storyworld evaluations
- The suspense algorithm which describes the encounter between reader, story and storyworld

The story can be considered the work of a hypothetical author, who first chooses some events from a storyworld and orders them into a fabula. The author then chooses events and orderings of events from this fabula to create a story which will hopefully trigger specific reactions for its readers. The evaluation of a given narrative thread corresponds to the reader's *static* evaluation of the state of the storyworld on completion of the narrative thread, and the suspense algorithm can be thought of as modelling the reader's *dynamic* reactions during the telling of the story.

4.1 The definition of a storyworld

change to definition...

A storyworld $\mathbb{W} = (\mathbb{E}, \mathbb{N}, \mathbb{D})$ is made up of the following elements:

- $\mathbb{E}$, the set of possible events, $\mathbb{E} = \{e_1, e_2, e_3, \ldots\}$[1],
- $\mathbb{N}$, the set of narrative threads. Each narrative thread $Z \in \mathbb{N}$ consists of a **fixed sequence** of distinct events chosen from the set $\mathbb{E}$ and an **importance value** $\text{Importance}(Z)$,
- $\mathbb{D}$, the set of ordered pairs (a, b) of disallowing events where $a, b \in \mathbb{E}$.

4.1.1 Chronology and the definitions of a story and a fabula

We will be dealing in this research with **chronological stories**. As narrative threads are sequences of events that typically occur chronologically, we use this very quality to define a chronological story in our system. A chronological story must comply with the *chronological constraint*.

For a given set of narrative threads, a story will satisfy the chronological constraint if and only if:

[1] In all the following, curly brackets ($\{\ldots\}$) are used for sets, that is, a number of *unordered* elements, square brackets ($[\ldots]$) are used for *ordered* sequences of elements and curved brackets ($(\ldots)$) for *ordered pairs*.

> For all pairs of events A and B where A precedes B in the story, if there *are* any threads in which both A and B occur, then in at least *one* of these threads A precedes B.

So a story can satisfy the chronological constraint even if there are no threads in which A and B occur and even if there is *some* thread in which B precedes A, as long as there is at least *one* thread in which A precedes B. This chronological constraint thus permits a certain 'looseness'.

We will allow for the possibility that some events in a narrative thread are skipped during the telling of a story, but will assume the existence of an underlying *fabula* in which no event is skipped. The only allowable difference in our model between *fabula* and *syuzhet* or, as we shall henceforth call it, *story*, is then that some elements of the *fabula* can be omitted in the *syuzhet*.

Using the chronological qualities of narrative threads, we can now give the definition of a **fabula** F as a chronologically ordered list of n events chosen from $\mathbb{E}$, the set of possible events in the storyworld $\mathbb{W}$:

Definition 1 *Fabula*

$$\begin{aligned}
&F \textit{ is a fabula for a storyworld } \mathbb{W} = (\mathbb{E}, \mathbb{N}, \mathbb{D})\\
&\qquad\qquad \Longleftrightarrow F = [e_1, e_2, \ldots, e_n] \textit{ where } n \geq 2,\\
&\qquad\qquad\quad \forall\ e_l \in F, e_l \in \mathbb{E}\\
&\qquad\qquad \textit{and if } M = \{T \in \mathbb{N} : e_l, e_m \in T\}\\
&\qquad\qquad\quad \forall\ l, m : 1 \leq l < m \leq n,\\
&\qquad \textit{then, either } M = \varnothing,\\
&\qquad\qquad \textit{or } \exists\ Z = [z_1, z_2, \ldots, z_k] \in M :\\
&\qquad\qquad\quad z_i = e_l, z_j = e_m\\
&\qquad\qquad\quad \textit{and } 1 \leq i < j \leq k
\end{aligned} \tag{4.1}$$

We define a **story** S as an ordered list of events *chosen* from a given fabula F. In general, a story for a fabula can reorder, repeat or skip any of the events in the fabula. We can express the general relation between the two by the following definition:

Definition 2 *Story*

$$\begin{aligned} S \text{ is a story for fabula } F \iff & S = [s_1, s_2, \ldots, s_k] \text{ where } k \geq 1, \text{ and} \\ & \forall\ s_i \in S, s_i \in F \end{aligned} \tag{4.2}$$

As we are restricting ourselves to *chronological* stories, we cannot change the *order* of events in the fabula to tell the story. Neither can we *repeat* fabula events to create a story, as we stipulated in Definition 4.1, the definition of a storyworld. The only possible variation between fabula and story is the omission of certain events. We can define chronological stories thus:

Definition 3 *Condition for a chronological story*

$$\begin{aligned} &\text{If } S = [s_1, s_2, \ldots, s_k] \text{ and } F = [e_1, e_2, \ldots, e_n], \text{ then} \\ &\quad S \text{ is a chronological story for fabula } F \iff S \text{ is a story for } F, \\ &\qquad \forall\ i, j : 1 \leq i < j \leq k, \\ &\qquad \exists\ l, m : 1 \leq l < m \leq n, \\ &\qquad s_i = e_l, s_j = e_m \end{aligned} \tag{4.3}$$

For example, referring to our fabula notation in Definition 1, the following is a possible story S for fabula F:

$$S = [e_1, e_5, e_6, e_9, \ldots, e_m]$$

4.1.2 Storyworld definitions

We now give some definitions concerning the interaction between fabulas and a storyworld $\mathbb{W} = (\mathbb{E}, \mathbb{N}, \mathbb{D})$. First we look at a possible constraint on the relation between $\mathbb{E}$ and $\mathbb{N}$.

A completeness relation between events and threads

The following (optional) completeness relation between the set of events, $\mathbb{E}$, and the set of narrative threads, $\mathbb{N}$, is a useful constraint to include in most storyworlds. It excludes the possibility that an event in a fabula has no narrative thread which contains it, thus avoiding the situation where an event is 'uninterpretable' in storyworld terms.

Definition 4 *Storyworld completeness*

$$\begin{aligned} \mathbb{W} = (\mathbb{E}, \mathbb{N}, \mathbb{D}) \textit{ is complete } \iff & \forall\ e \in \mathbb{E}, \\ & \exists\ Z \in \mathbb{N} : e \in Z \end{aligned} \tag{4.4}$$

Now we look at the role of $\mathbb{D}$, the set of disallowing event-pairs.

Disallowing event-pair definitions

Informally, if $(a, b) \in \mathbb{D}$, the set of disallowing pairs, this means that if a is told, then b is predicted *not* to occur in storyworld $\mathbb{W}$, or we can also say, b *should not* be one of the subsequent events to be told. We can express this in the following way:

Definition 5 *Disallowing event-pair conditions*

$$
\begin{aligned}
(a,b) \in \mathbb{D} \iff & \textit{The occurrence of event } a \textit{ in storyworld } \mathbb{W} \\
& \textit{has a physical and/or intentional causal effect} \\
& \textit{which renders impossible the future occurrence of event } b
\end{aligned}
\tag{4.5}
$$

Of course, storyworlds are different from real worlds in that they serve purely narrative purposes, and it is always theoretically possible to *tell* events which are in contradiction with each other. We can avoid this possibility by using only fabulas that are *(storyworld) consistent*. We can use the set $\mathbb{D}$ to give a definition for a consistent fabula:

Definition 6 *Condition for a consistent fabula*

$$
\begin{aligned}
F = [e_1, e_2, e_3, \ldots, e_n] \textit{ is a consistent fabula} \iff & \nexists\, i, j : i < j, \\
& (e_i, e_j) \in \mathbb{D}
\end{aligned}
\tag{4.6}
$$

In other words, we require that no event in the fabula disallows any other[2].

The ordered pair (a, b) defines the general case where a disallows b. However, in many cases (and in many storyworlds), if a disallows b, then the reverse is also true. For such cases, it is useful to define $\mathbb{D}_{mutual}$, a subset of $\mathbb{D}$, as follows:

Definition 7 *Mutually disallowing events*

$$
(a,b) \in \mathbb{D}_{mutual} \iff (a,b) \in \mathbb{D} \textit{ and } (b,a) \in \mathbb{D} \tag{4.7}
$$

[2]This is a transposition of the constraints in the GLAIVE system that we discussed earlier in 3.2.5.

4.2 Telling the story

4.2.1 Story states

Telling the story is equivalent to going through the ordered list of events in S one by one. During the telling of the story S, we maintain two ordered lists: told events, T, and untold events, U. We use an ordered pair which represents how much of story S has been told:

Definition 8 *Representing the state of the story*

$$\text{State}(S) = (T, U) \tag{4.8}$$

If the story contains m events, we can define an index n from *1* to m, to describe the n^{th} event in the story. We then have $\text{State}_n(S)$ as the state of the story S after event n has been told. To 'tell an event in the story', we take the next untold event from the list of untold events U and add it to the tail of the list of told events T. We can formulate the updating of S_n as we go through each story event as follows:

Definition 9 *Telling the story*

$$\begin{aligned} &\textit{For a story } S, \\ &\textit{if } \text{State}_n(S) = ([s_1, s_2, \ldots, s_n], [s_{n+1}, s_{n+2}, s_{n+3}, \ldots, s_m]) \\ \textit{then } &\text{State}_{n+1}(S) = ([s_1, s_2, \ldots, s_n, s_{n+1}], [s_{n+2}, s_{n+3}, \ldots, s_m]) \end{aligned} \tag{4.9}$$

4.2.2 Narrative thread states

Narrative threads can be *active* or *inactive*. As the story S is told, threads may change from being active to inactive and vice versa.

In a similar definition to that given for the story S, events in a thread can be shifted from one list into another. This can occur either as events are *told* as part of the story, or as they are *implicated* as part of a sequence of events. Told events are *known* to have occurred in the storyworld, whereas implicated events are only assumed to have *potentially* occurred. We will use the term *convey* to group together both of these cases. Etymologically, the term comes from the Latin *con-viare* which means: 'travelling with'. It also has associations with a conveyor belt, where individual items are carried along by a general movement.

For each thread Z, therefore, we designate state(Z) which indicates both whether Z is *active* or *inactive* and which events in Z have been *Conveyed* (C), and which are as yet *Unconveyed* (U). We notate this in the following way:

Definition 10 *Representing the state of a narrative thread*

$$\text{state}(Z) = (\mathit{active}|\mathit{inactive}, C, U) \tag{4.10}$$

We now describe the syntactic rules which our narrative thread system follows to model suspense in narrative.

Rule 1 *Initial conditions*

> *Before the story starts to be told, all the narrative threads $Z \in \mathbb{N}$ have the following state:*
>
> $$\text{state}(Z) = (\text{inactive}, \emptyset, U) \tag{4.11}$$

This means that their state is *inactive* and that they have *no* conveyed events, or alternatively, all their events are *unconveyed. Inactive* threads always have this form and have no effect on suspense calculations.

Rule 2 *Thread activation*

> *When the q^{th} event, α_q, of the m events of thread Z is* told *in the story, we have:*

$$\text{state}(Z) = (\text{active}, [\alpha_1, \alpha_2, \ldots, \alpha_q], [\alpha_{q+1}, \alpha_{q+2}, \ldots, \alpha_m]) \tag{4.12}$$

In this way, all events in Z which precede α_q *also* get placed in the list of conveyed events of Z. This mechanism allows for **ellipsis** in narrative: some events can be treated *like* told events even though they have not been told in the story, just because they precede a told event in some narrative thread.

Rule 3 *Thread success*

> *When the last event in a narrative thread Z gets conveyed in the story, we can say: 'Z succeeds' and we have the following:*

$$\text{state}(Z) = (\text{active}, C, \emptyset) \tag{4.13}$$

> *Just before the very next event in the story is told, such a narrative thread becomes* inactive *as follows:*

$$\text{state}(Z) = (\text{inactive}, C, \emptyset) \tag{4.14}$$

Before describing our next rule, we need to give some further definitions for events and threads.

Event definitions

Definition 11 *Implicated prior events*

> *An implicated prior event is any event in the Conveyed list of some active narrative thread that has* not *been told in the story.*

The *adding with ellipsis* mechanism described above leads to situations where some of the events included in the *Conveyed* list of a given thread will *not* have been told in the story. We call these events *implicated prior events*. Unlike told events, implicated prior events are defeasible, that is, later story events may reveal that they did *not* actually occur in the storyworld by disallowing the thread containing them. However, insofar as a thread containing such events is held to be a correct interpretation of the underlying fabula of the story currently being told, then these events will very often simply be assumed to have occurred in the storyworld. This is the reason for their inclusion in what we have called the *Conveyed* list of a thread.

Definition 12 *Conflicted implicated prior events*

> *If α and γ are implicated prior events (in different threads), and $(\gamma, \alpha) \in \mathbb{D}$, then α is a conflicted implicated prior event.*

In a similar way to implicated *upcoming* events, implicated prior events in different threads may remain in conflict with each other over several story steps. Conflicted implicated prior events are important for our concept of *revelatory suspense* which we discuss in 4.3.

Definition 13 *Implicated upcoming events*

> *An implicated upcoming event is any event that is a member of the* Unconveyed *list of some active thread.*

Such an event is predicted to be told in the story being told with a confidence level that depends on the confidence we have in the narrative thread of which it is a member. It is conflicts between implicated upcoming events that create suspense.

Thread definitions

Definition 14 *Confirmed threads*

> *A confirmed thread is an active thread whose* Conveyed *list contains at least onetold event.*

It follows that *unconfirmed* threads contain no event which has yet been told in the story. Such threads are mostly *inactive*, but may become *active* while staying unconfirmed in certain conditions as we shall see.

Definition 15 *Disallowed threads*

> *An active narrative thread with an event α in its Unconveyed list is disallowed when an event γ is told in the story and $(\gamma, \alpha) \in \mathbb{D}$.*

Such a thread becomes *inactive* and can no longer become *active*.

Definition 16 *Conflicted threads*

> *A conflicted thread is a thread whose Conveyed list contains at least one conflicted implicated prior event.*

Such a conflicted prior event, α, will be in conflict with another implicated prior event γ in a different active narrative thread, because $(\gamma, \alpha) \in \mathbb{D}$. There will necessarily be a degree of uncertainty about whether a conflicted thread Z is the correct interpretation of story events because there is at least one other active thread that contains an event which *would have disallowed* Z had it been told in the story and not just implicated through the *adding with ellipsis* rule. Conflicted threads are important for *revelatory suspense* (discussed in 4.3).

Definition 17 *Active unconfirmed threads*

> *An inactive thread Z can become an* active unconfirmed thread *if any of its (unconveyed) events appears in the Unconveyed list of some* active *confirmed thread (and as long as it has no event that is disallowed by some told event).*

Thus, in such a case, an inactive (and thus unconfirmed) thread becomes active even though none of its events have yet been told in the story. We can say that 'the confirmation of thread Z is predicted'.

Such threads have the general form state$(active, \emptyset, U)$ and can have an effect on suspense. We can formulate their activation by the following rule:

Rule 4 *Activation of unconfirmed threads*

$$\begin{aligned} &\textit{if, for threads } Z, Y, \textit{ we have} \\ &\qquad \text{state}(Z) = (\textit{inactive}, \emptyset, U_z), \\ &\qquad \text{state}(Y) = (\textit{active}, C_y, U_y), \\ &\qquad \textit{and } \exists\ \alpha : \alpha \in U_y, \alpha \in U_z, \\ &\textit{then set } \text{state}(Z) = (\textit{active}, \emptyset, U_z) \end{aligned} \tag{4.15}$$

Active unconfirmed threads are important in our system because they allow for some degree of flexibility in the linking together of narrative chains. This rule in effect allows two threads which share at least one event to function together for the purposes of suspense calculation. For example, a set of threads which detail the different things that someone might do when they get home can under this rule be activated before the story narrates the moment when they open their front door.

4.2.3 A global story-telling variable

We define $\mathrm{Threads}_n(\mathbb{N})$, the set of all narrative thread states after the n^{th} event has been told, as follows:

Definition 18 *The set of all narrative thread states*

$$\mathrm{Threads}_n(\mathbb{N}) = \{\mathrm{state}_n(Z) : Z \in \mathbb{N}\} \tag{4.16}$$

Now we can define a global variable $\mathrm{G}_n(S, \mathbb{N})$ which represents the current state of the story-telling process.

Definition 19 *The global state of the story-telling process*

$$\mathrm{G}_n(S, \mathbb{N}) = (\mathrm{State}_n(S), \mathrm{Threads}_n(\mathbb{N})) \tag{4.17}$$

> *where* $\mathrm{State}_n(S)$ *represents the state of the story* S *and* $\mathrm{Threads}_n(\mathbb{N})$ *represents the states of all the narrative threads in the story-world* $\mathbb{W}$ *after event* n *has been told.*

4.3 Modelling revelatory suspense

We now clarify some of the rules that allow us to model revelatory suspense by using partially confirmed events.

A given story event may be present in several different threads. When a partially unexplained or strange event δ is told in the story, we model this by assuming that *several* threads become activated as candidates to become *the* thread which definitively includes (and thus explains) δ. Subsequent events in the story may disallow some of these candidate threads. Exactly which thread turns out to be the 'correct interpretation' of δ is determined by the rest of the story.

A simple example will make this clearer. Suppose we have the following narrative threads:

$$\begin{aligned} T &= [e_0, e_2, e_4] \in \mathbb{N} \\ Z &= [e_1, e_2, e_3] \in \mathbb{N} \end{aligned} \tag{4.18}$$

Suppose in addition we have the following mutually disallowing pairs:

$$(e_0, e_1) \in \mathbb{D}_{mutual}, (e_3, e_4) \in \mathbb{D}_{mutual} \tag{4.19}$$

These disallowing pairs indicate that in this very basic storyworld there are only two possible fabulas (or subsets thereof) which can be told:

$$\begin{aligned} F_T &= [e_0, e_2, e_4], \\ F_Z &= [e_1, e_2, e_3] \end{aligned} \tag{4.20}$$

Now suppose our story starts with event e_2. This event belongs to both fabulas so we cannot yet determine which one is being told. In other words, we do not yet know which thread corresponds to the fabula that is being told[3].

Once event e_2 has been told and the threads have been updated, we end up with the following narrative thread states:

$$\begin{aligned} \text{state}(T) &= (active, [e_0, e_2], [e_4]) \\ \text{state}(Z) &= (active, [e_1, e_2], [e_3]) \end{aligned} \tag{4.21}$$

where events e_0 and e_1 have been moved to the *Conveyed* list of their respective threads thanks to the *adding with ellipsis* rule. Notice however

[3]Of course, as soon as the next event gets told, e_3 or e_4, one of the threads T or Z will be disallowed and the matter will be settled.

that e_0 and e_1 are incompatible, that is, they would have disallowed each other, if either of them had been told in the story. One of these two events must belong to the *fabula* being told even though it has been omitted in this particular *story*. As we have seen, we call such events *implicated prior events*. Such events may conflict with each other and yet still not disallow each other's threads. In fact, this is a rule of our system:

Rule 5 *Told events rule*

> *Only events that are told in the story can disallow threads.*

The fact that we have conflicted events that precede the current event is an example of the uncertainty of interpretation that creates revelatory suspense. To give an example of this, we show Figure 4.1 where narrative threads A and B share a common event, the event which has just been told in the story. We can see that thread A has two implicated prior events which are in conflict with other implicated prior events from thread B. At this point in the story, thread A thus has *two* conflicted prior events and *one* confirmed event.

The number of *conflicted prior events* for a given thread will give us a negatively correlated measure of the confidence with which this thread is held to be valid description of the story thus far. If there are *many* conflicted prior events, then we will have *low* confidence in the thread.

On the other hand, the number of *told* events for a given thread will give us a positively correlated measure of the confidence we have for the thread. If there are many *told* events, then we will have *high* confidence.

Combining these two opposing confidence measures, we can extrapolate a single confidence measure for all threads. Threads will thus have varying

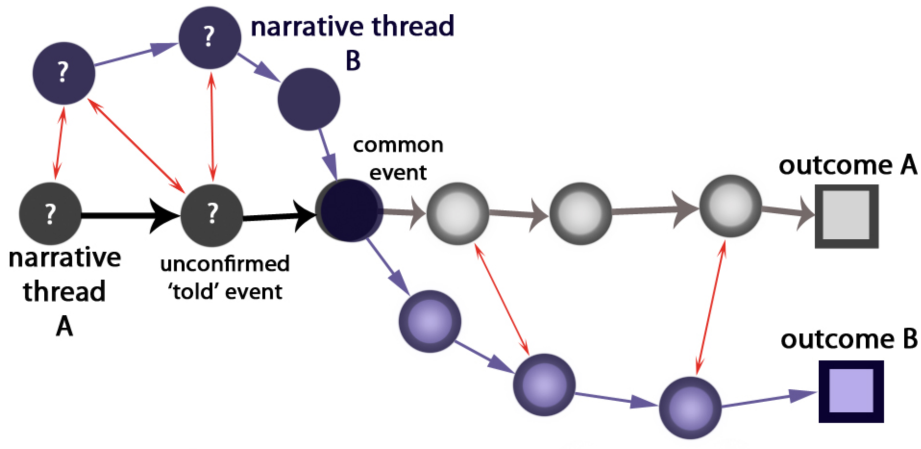


Figure 4.1: Implicated prior events creating revelatory suspense

Filled-in shapes represent conveyed events:

- filled-in shapes without a question mark have been *told* in the story
- shapes with a question mark represent *implicated prior events*

Empty shapes represent events that have not *yet* been told.

degrees of *confidence* as the story progresses. This fluctuating confidence level is a ongoing revelatory process which actually creates its own type of suspense. A given narrative thread Z will thus have its potential effect on suspense reduced, if it has many *conflicted* prior events. Note that as the other threads that contain the events which conflict with Z get disallowed or become inactive, then Z may come to no longer have any *conflicted* prior events. In this case, if Z has at least *one* confirmed event, its confidence level will reach the maximum value of 1.

To satisfy the above relation and boundary conditions, we used the following formula as a measure for the Confidence of a narrative thread:

Definition 20 *Confidence*

> *If P is the (non-zero) number of confirmed events and Q the number of conflicted prior events of an active thread Z, then the Confidence(Z) with which thread Z is considered a valid*

interpretation of the events in the story is defined as follows:

$$\begin{aligned} \text{Confidence}(Z) &= \frac{1}{(1+\frac{\phi Q}{P})} \\ &= \frac{P}{(P+\phi Q)} \end{aligned} \tag{4.22}$$

for some ***conflicted-to-told ratio*** $\phi : 0 < \phi < \infty$.

For the case $P = 0$*, that is, for any active unconfirmed thread* Z*, Confidence*(Z) *is defined as being equal to the Confidence value of the thread that triggered the activation of* Z*. Such an unconfirmed thread may later get confirmed in which case* $P \neq 0$ *and its Confidence value can be calculated as above.*

In our model, we used $\phi = \mathbf{1.5}$. The higher this number, the more we boost the effect of conflicted prior events over already told events. For a high number of conflicted events, Q, the Confidence will tend to zero. For a high number of told events, P, the Confidence will tend to 1. We therefore have $0 < \mathit{Confidence} \leq 1$.

4.4 Modelling the reader's predicted reactions

4.4.1 Importance values

To each narrative thread Z, we associate an **Importance value** Importance(Z), which can be positive or negative.

Definition 21 *Importance value* Importance(Z)

Each thread Z *has* Importance(Z) *where*

Importance(Z) = *the predicted degree of positive or negative appraisal of the storyworld situation that the reader would have, were* Z *to succeed.*

We assume that all the affinities the reader has with the events in the storyworld are known. This means that once the final event of a given thread has been told, we can always determine a clear-cut Importance value for the thread in question when we evaluate the state of the storyworld. Typically, this value would be positive if the event is good for the hero and negative if it is good for the villain.

In our model, we will use the range $(-10, +10)$ for the Importance value, where $+10$ corresponds to an event about which the reader is very positive (happy, overjoyed, satisfied) and -10 corresponds to an event about which the reader is very negative (sad, gloomy, dissatisfied).

4.4.2 Potentially useful simplifying assumptions

The following definitions describe some potentially useful assumptions that can be made about the Importance variable. They can be added to the general mathematical model, but are not essential to it.

Definition 22 *Importance values and thread success*

> *The Importance value* Importance(Z) *of a narrative thread* Z *is defined* only *in relation to the state of the storyworld after its* final *event has been told.*

For example, suppose we have the following narrative thread Z:

$$Z = [\text{gets arrested}, \text{gets tried in court}, \text{gets convicted}, \text{gets sent to jail}]$$

In our model, such a thread could only be considered to be bad for a story character because of the state of the storyworld after the final event: '*gets sent to jail*'. The ascription of the Importance value for this thread would therefore be based uniquely on the appraisal of the storyworld once the thread has succeeded, that is, once the event '*gets sent to jail*' has been told.

Under this assumption and referring to narrative thread Z above, we would therefore be unable to ascribe an Importance value to, say, the event: '*gets arrested*' (unless of course '*gets arrested*' were to be the final event in a *different* thread). However, the event '*gets arrested*' may stop other positive events in the storyworld from occurring, for example: '*walks home*', '*relaxes at home*', and so on. So, under this assumption, a *non-final event* such as '*gets arrested*' could have a negative effect on reader expectations for the main character and provoke changes in the perceived suspense level in only two possible ways: i) by potentially disallowing positive events, and ii) by being part of a thread whose success results in a negative storyworld appraisal. It could not have its own specific imminence-related suspense effect.

Definition 23 *Constancy of Importance values*

> *The Importance values of all threads remain* constant *whatever happens in the story.*

This optional assumption creates the constraint that the Importance value of a thread cannot be changed by the success or failure of other threads. The reader's evaluation of what good or bad events are in the storyworld would remain constant during the telling of the story. If this evaluation were to be based on the success or failure of a character in the story, for example, then the importance to the reader of this character's success or failure would not vary during the telling of the story.

Of course, in general such an assumption will not hold over the whole duration of a even moderately long story, and that *re-evaluations* of the importance of potential events in a storyworld are part and parcel of what stories try to achieve. In the later implementation of our mathematical model described in chapter 5, however, we will be using very short stories which will make this assumption both useful and plausible. Importantly, by not having to cater for the complexities of longer, more involved stories, we will be able to more easily concentrate on the *fundamental* mechanisms of suspense modelling.

4.5 The general algorithm for suspense

We now describe in detail a possible general algorithm for evaluating the suspense level after the telling of each new event of a story S in a storyworld $\mathbb{W}$.

4.5.1 The story-telling update process

The new narrative thread states, $\text{Threads}_{n+1}(N)$ are a function of the old ones, $\text{Threads}_n(N)$, together with the latest newly told $(n+1)^{th}$ event in the story which we will call δ:

$$\text{Threads}_{n+1}(N) = \mathit{update}(\delta, \text{Threads}_n(N)) \tag{4.23}$$

This *update* function consists of three basic procedures which modify the state of each narrative thread in N. These are performed in the following order:

- Adding threads
- Equalising threads
- Disallowing threads

- Completing threads

Adding (with ellipsis)

If the new story event δ matches the k^{th} member of the *Unconveyed* list of any narrative thread Z, then we move it (and all the events before it) into the *Conveyed* list of the thread. Additionally, the thread also becomes active (if it was not previously). Mathematically, we can write the following:

Rule 6 *Adding*

$$
\begin{gathered}
\textit{if } \delta \in S \quad \textit{is the } (n+1)^{th} \textit{ story event,} \\
\textit{then, for all threads } Z \textit{ such that} \\
\mathrm{state}_n(Z) = (\textit{active}|\textit{inactive},[t_1, t_2, \ldots, t_m], [u_0, u_1, \ldots, \delta, u_{k+1}, u_{k+2}, \ldots u_{final}]), \\
\mathrm{state}_{n+1}(Z) = (\textit{active},[t_1, t_2, \ldots, t_m, u_0, u_1, \ldots, \delta], [u_{k+1}, u_{k+2}, \ldots u_{final}])
\end{gathered}
\tag{4.24}
$$

Note that under this algorithm, all events preceding the new event δ are presumed to have occurred in the storyworld and are moved into the set of *Conveyed* events, even though they may not have been told in the story. We label this the **adding with ellipsis rule** and it accounts for the fact that if an event δ occurs in the storyworld, then the events for which it is a typical sequel may well have also occurred, for as we know, stories contain ellipses.

Equalising

This step is needed for consistency. It ensures that all the *implicated* events that are moved from the Unconveyed to the Conveyed list of a given narrative thread Z due to Rule 6 above, also get moved from the Unconveyed to the Conveyed list of all other threads in a similar way.

Rule 7 *Equalising*

$$\begin{aligned}
&\forall\, Z : \mathrm{state}(Z) = (\mathit{active}, C_z, U_z),\\
&\forall\, Y : \mathrm{state}(Y) = (\mathit{active} \mid \mathit{inactive}, C_y, U_y),\\
&\qquad \mathit{if}\ \exists\,\gamma : \gamma \in C_z, \gamma \in U_y \qquad (4.25)\\
&\mathit{then\ shift\ all\ events\ in}\ U_y\ \mathit{up\ to\ and\ including}\ \gamma,\\
&\qquad\qquad \mathit{into\ list}\ C_y
\end{aligned}$$

Threads which change their state in this way get their *Foregroundedness* set to *1* and also become *active* if they were not before.

Disallowing

Next we must deactivate all the threads that may have been disallowed by the new story event[4]:

Rule 8 *Disallowing*

$$\begin{aligned}
&\mathit{if}\ \delta \in S\ \ \mathit{is\ the}\ (n+1)^{th}\ \mathit{story\ event,\ and}\\
&\quad (\delta, \gamma) \in \mathbb{D},\ \mathit{the\ set\ of\ disallowing\ event\text{-}pairs},\\
&\mathit{then,\ for\ all\ threads}\ Z\ \mathit{such\ that}\\
&\quad \mathrm{state}_n(Z) = (\mathit{active}, \mathit{Conveyed}, \mathit{Unconveyed})\ \mathit{and}\ \gamma \in \mathit{Unconveyed},\\
&\quad \mathrm{state}_{n+1}(Z) = (\mathit{inactive}, \emptyset, \mathit{Conveyed}\ + \mathit{Unconveyed})
\end{aligned} \qquad (4.26)$$

[4]Note that in this example, if $\gamma \in \mathit{Conveyed}$, that is, if an event is told in the story which disallows an event that has already been told, then we leave the narrative thread in question unchanged. We assume that an already told event cannot be revoked, and nor can the thread that contains it, when disallowing information arrives ‘too late’.

Completing

During these first two phases, any thread whose final event has been Conveyed remains active in order to allow the calculation of all its interaction with the other threads. Once the disallowing procedure has been carried out, the thread is considered to have succeeded and its status reverts to being inactive. Mathematically, we combine the two equations 4.13 and 4.14 to give the following rule:

Rule 9 *Completing*

$$\begin{aligned} \forall\, Z : \mathrm{state}_n(Z) &= (\mathit{active}, [t_1, t_2, \ldots, t_m], \emptyset), \\ \mathit{set}\ \ \mathrm{state}_{n+1}(Z) &= (\mathit{inactive}, [t_1, t_2, \ldots, t_m], \emptyset) \end{aligned} \tag{4.27}$$

4.5.2 Suspense evaluation

Once we have updated the states of the narrative threads, we can apply the following algorithm which produces a measure of the suspense level of the story after each event. Following our preceding analysis, we first determine the following intermediate values :

- Imminence
- Importance
- Foregroundedness
- Confidence

We then combine these values to produce a suspense value for each *individual* narrative thread. Finally we suggest a heuristic which combines all the individual narrative thread suspense values to produce the *global* suspense level at any moment in a story.

Imminence

Each active narrative thread Z generates two values for Imminence:

- Completion Imminence: this is related to the number of events in Z still to be conveyed for it to be completed or to 'succeed'.
- Interruption Imminence: this is related to the smallest number of events still to be conveyed in some other thread Y before an event is told which can interrupt Z by disallowing one of its events. In the case where no thread can interrupt Z, the Interruption Imminence of Z is defined as zero.

Every active narrative thread will therefore produce suspense both due to its potential *completion* and due to its potential *interruption* by disallowing events in other threads. Inactive threads produce a suspense value of zero.

Completion Imminence We set the Completion Imminence Number H for an *uncompleted* thread Z as equivalent to the number of events left to be conveyed in the thread[5]. Mathematically, this gives the following:

Definition 24 *Completion Imminence Number H of a thread*

$$\begin{aligned} &\forall\, Z \in N,\ \textit{where}\ \mathrm{state}_n(Z) = (\textit{active}, T, U), U \neq \emptyset \\ &\qquad\qquad\qquad\qquad H = \ |U|,\ \textit{the number of elements in } U \end{aligned} \tag{4.28}$$

[5]In fact, this definition of Completion Imminence was used only for *confirmed* active threads. The procedure to obtain the Completion Imminence number for *unconfirmed* active threads is detailed in 5.1.1. A different method of calculation was used to take into account the fact that unconfirmed threads may be dependent on events that happen much later in a given story than the current event in order to be confirmed at all. The completion of such threads cannot therefore be seen as having the same imminence as confirmed threads.

This variable has the following range: $1 \leq H < \infty$. It cannot be zero, for this would mean the narrative thread would be completed and in this case the Imminence level is *defined* as zero.

To illustrate Completion Imminence, we show a thread with a Completion Imminence number of 4 in Figure 4.2.

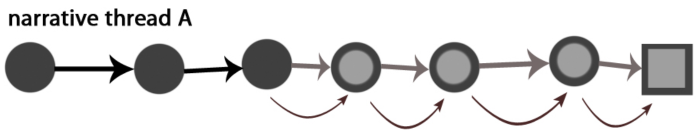


Figure 4.2: Completion Imminence

Interruption Imminence The **Interruption Imminence** number R of a thread Z is related in our model to the number of *unconveyed* events left in another active thread Y that need to be told before one of these can disallow an unconveyed event in Z. We will make the assumption that a thread must be *active* in order to potentially interrupt another[6]. Not all threads are interruptible by all threads so a given thread Z may have zero, one or many potentially interrupting threads. To deal with these cases, we first define $R(Z, Y)$, the interruption number R of thread Z by thread Y as follows:

[6]Again, the following definition of Interruption Imminence was used only for *confirmed* active threads. For the same reasons as in the case of Completion Imminence above, the exact procedure to obtain the Interruption Imminence number for *unconfirmed* active threads differs and is detailed in 5.1.1.

Definition 25 *Interruption Imminence Number $R(Z,Y)$*

$$\begin{aligned} \forall Z,Y \in N,\ \textit{where}\ \mathrm{state}_n(Z) = &\ (\textit{active}|\textit{inactive}, C_z, U_z), \\ \mathrm{state}_n(Y) = &\ (\textit{active}, C_y, U_y), \\ \textit{then},\ \forall\,(u_y,u_z) \in \mathbb{D}, u_y \in U_y\ &\textit{and}\ u_z \in U_z, \\ \textit{then if}\ m = \min_{\forall y,z}\{\ &\textit{number of elements in}\ U_y\ \textit{before}\ u_y\} \\ R(Z,Y) = &\ m+1, \\ \textit{otherwise}\ R(Z,Y) = &\ \infty \end{aligned} \tag{4.29}$$

Similarly, this variable also has the following range: $1 \leq R < \infty$.

To illustrate the Interruption Imminence between two threads, in Figure 4.3, we show thread A which has an Interruption Imminence number of 3 due to thread B.

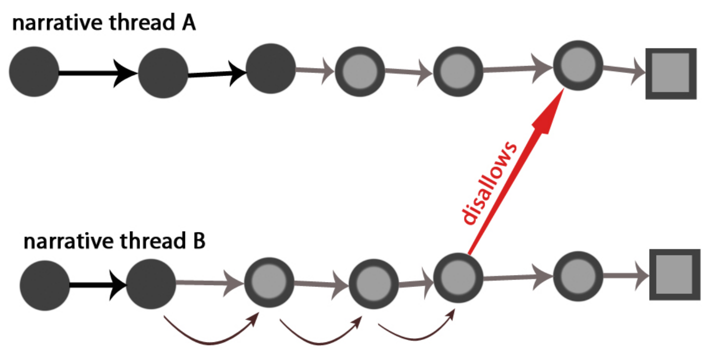


Figure 4.3: Interruption Imminence

We can now deal with the general case where a thread could be interrupted by several threads. From all the individual interruption imminence values that each interrupting thread produces, we pick out the single value of the interrupting thread with the *highest* imminence, or in other words, the *lowest*

Interruption Imminence Number. If there are no interrupting threads, this value will be infinite.

Definition 26 *Interruption Imminence Number $R(Z)$ for thread Z*

$$R(Z) = \min_{\forall Y \in N}(R(Z, Y)) \tag{4.30}$$

Total Imminence We can now give a first definition of Total Imminence$_n(Z)$, a measure of the Total Imminence of a narrative thread Z after the n^{th} event in the story. To enable exploration of the relative effects of Completion Imminence and Interruption Imminence on this measure, we create a factor ρ which can vary and weight these two effects. This leads to the following complete formula for the total imminence:

Definition 27 *Total Imminence: the general case*

$$\begin{aligned} \textit{Total Imminence}_n(Z) = {} & \rho.\textit{imminenceFunction}(H_n) \\ & + (1-\rho).\textit{imminenceFunction}(R_n), \end{aligned} \tag{4.31}$$

for some imminenceFunction to be defined, where H_n is the Completion Imminence number, R_n is the Interruption Imminence number for thread Z after the n^{th} event in the story and ρ is a weighting factor.

If ρ were set to 0.5, then the relative effect of Completion Imminence and Interruption Imminence would be the same. The results from the implementation of our model which we describe later led us to choose $\rho = 0.7$, thus boosting the effect of *Completion imminence.*

If a *large* number of events must be told for a thread to be completed, the Imminence is *low*, and vice versa. A simple Imminence function could therefore take one of the following forms: $1/x$ or e^{-x}. We adopted a variant of the first of these two options, leading to the following definition of an Imminence value for a given thread Z (where we have omitted the n subscripts)[7]:

Definition 28 *Total Imminence of thread Z*

$$\textit{Total Imminence} = \rho\frac{1}{H} + (1-\rho)\frac{1}{R} \tag{4.32}$$

> *where $\rho = 0.7$, H is the number of events to the completion of Z and R is the minimum number of events before an event in some other narrative thread could be told which would disallow some unconveyed event in Z.*

Importance

We define Importance$_n(Z)$, a measure of the potentially variable Importance of a narrative thread Z after the n^{th} event of the story has been told. In our model, the Importance value of a thread is just a scalar factor which boosts or weakens its effect on perceived suspense compared to all other threads. Our model uses the range $(-10, +10)$ for the Importance value of a thread. In the implementation of our model described in chapter 5, the Importance$_n(Z)$ is taken to be a constant and thus does not vary with n.

[7] There could be an argument for making Imminence depend proportionally to the *ratio* of conveyed and unconveyed events in a thread. However, if the temporal characteristics of narrative events are the same for all threads, then it does not seem logical to make the Imminence depend on the length of a particular thread: an outcome one event away is one event away for all thread lengths. However, if there were a idea of a standard length of narrative thread, say 7 events, then threads which would be shorter than this, say a 3-event thread could be considered to be approximations for the full 7-event version. In this case, it would be appropriate to consider using the proportion of remaining, unconveyed events compared to the total number of events. For our 3-event example, if 2 events had been conveyed out of 3 then there would be a projected number of actual unconveyed events of $1/3 \times 7 = 7/3$ events and thus a Completion Imminence of $1/(7/3) = 0.43$ a significantly lower value than the current $1/(1) = 1$. But the adoption of such a model for Imminence depends wholly on this unproven, albeit interesting idea of a standard 'underlying' length of a narrative thread.

Foregroundedness

We use the term **Foregroundedness** as a function of how present a given narrative thread is in the reader's mind. The level of Foregroundedness ascribed to each active narrative thread changes with each new event that is told in the story. We set this parameter to vary between 0 and 1.

Any narrative threads which contain the current story event are considered to be very present in the mind of the reader and get ascribed the maximum level of Foregroundedness, that is 1. The Foregroundedness of all *other* narrative threads decreases at each story step.

We can thus define $\text{Foregroundedness}_n(Z)$, a measure of the Foregroundedness of a narrative thread Z after the n^{th} event in the story. The Foregroundedness of all threads is recalculated after each new story event as follows[8]:

Definition 29 *Foregroundedness*

$$\begin{aligned}
&\textit{if } \delta \in S \textit{ is the } (n+1)^{th} \textit{story event,} \\
&\textit{then, for all threads } Z \textit{ such that} \\
&\qquad \text{state}_{n+1}(Z) = (\textit{active}, C, U) \textit{ and } \delta \in C, \\
&\qquad \textit{Foregroundedness}_{n+1}(Z) = 1
\end{aligned} \tag{4.33}$$

Narrative threads which do *not* contain the current story event undergo a decrease in foregroundedness due to a decay function in the following way:

$$\text{Foregroundedness}_{n+1}(Z) = \text{decayFunction}(\text{Foregroundedness}_n(Z)) \tag{4.34}$$

[8]With this formulation, the Foregroundedness will be set to 1 when the new event δ is any member of the list of *Conveyed* events, that is when $\delta \in T$, and not just the most recently conveyed event. This covers the case when an event is retold or re-mentioned in some way. We consider that the thread also becomes foregrounded for the reader in such cases.

A simple decay function could take the following form:

$$\text{decayFunction}(x) = \beta x, \text{where } 0 < \beta < 1 \tag{4.35}$$

Experimentation led us to use $\beta = 0.88$. We also have $0 < \text{Foregroundedness} \leq 1$.

Confidence

Following the derivation that we presented in 4.3, we also include $\text{Confidence}_n(Z)$ in our suspense evaluation, a measure of the Confidence of a narrative thread Z after the n^{th} event in the story.

$$\text{Confidence}_n(Z) = \frac{1}{(1 + \frac{\phi Q}{P})} = \frac{P}{(P + \phi Q)} \text{where } \phi = 1.5 \tag{4.36}$$

where P is the (non-zero) number of **told events** and Q the number of **conflicted implicatedprior events** of narrative thread Z. Empirical work on the implementation of our mathematical model to a domain led us to determine a conflicted-to-told ratio $\phi = 1.5$.

Combining suspense factors

After the telling of the n^{th} story event, for each active narrative thread Z, we calculate the $\textit{Imminence}_n(Z)$, $\textit{Foregroundedness}_n(Z)$, $\textit{Confidence}_n(Z)$ and $\textit{Importance}_n(Z)$. For the general case, we assume that all four variables are orthogonal to each other. We can therefore choose to calculate the contribution of each active narrative thread Z to the global suspense as the result of the *multiplication* of these values, as follows:

Definition 30 *Suspense contribution of thread Z*

$$\begin{aligned} \mathit{Suspense}_n(Z) = \ & \mathit{Imminence}_n(Z) \\ & \times \mathit{Importance}_n(Z) \\ & \times \mathit{Foregroundedness}_n(Z) \\ & \times \mathit{Confidence}_n(Z) \end{aligned} \tag{4.37}$$

There are other possible ways to combine these values to obtain a suspense value for a narrative thread. We use multiplication because our concern is to find the simplest possible model of suspense.

The global suspense at a given moment in a story

We can calculate the suspense contribution from each active narrative thread as above. But apart from what is at stake for each thread, we also have to consider what is at stake in the story as a whole.

In a given storyworld, it could be the case that different groups of threads represent different sub-stories within the overall story. Indeed, it may be possible to use the number of shared events between sets of narrative threads as a criterion for determining the presence of sub-stories in a narrative. In such a case each group of interlinked threads could contribute separately, *as a group*, to the global suspense level of the story. In this first approach to suspense modelling, however, we will only be considering *stories with no sub-story*, that is, where all threads are connected in some way.

In general, we can derive a value for the *Global suspense* from the set of suspense values of individual threads:

$$\text{Global suspense} = \mathit{globalSuspense}(\{\text{Suspense}(Z)_{\forall Z \in N}\})$$

There are many possible ways to define the function *globalSuspense*. We list here a few possibilities, all of which could be used and tested empirically:

1. We take the absolute value of the *one most extreme suspense value*, positive or negative, from all the current values.

2. We take the sum of the absolute values of the *two* most extreme suspense values, be they positive or negative.

3. We take the sum of the absolute values of *all* suspense values. An alternative would be to take the sum of the square of all suspense values. This approach uses both the *number and strengths* of all suspense values.

4. We take the *difference* between the highest positive value and the lowest negative value. This approach gives a measure of the *spread* in the suspense values. It might also be seen as a measure of 'what's at stake' at this moment in the story.

5. We use the *standard deviation* of all suspense values. This approach is based on the *relative dispersion* of the suspense values.

The approach we took in this research was to treat the global suspense value in a similar way to the way we treated each individual thread. As we discussed in 3.3.3, for any given active thread, we will eventually end up with one of two outcomes during the telling of the story: either the thread succeeds, in which case an effect equivalent to the Importance of the thread is produced, or the thread gets interrupted (or its foregroundedness becomes so low that its effect on suspense is nullified), in which case no effect is produced. In terms of a thread's effect, what-is-at-stake is simply equivalent to the difference between these two values and is equal to the thread's Importance.

In a similar way, we chose to adopt option 4. above and define the global suspense as the *difference between the best and worst possible outcomes*. In this case, a simple way to get a global suspense value is to just consider the highest and lowest suspense values taken from all the threads. This is a kind of first-past-the-post system, where only the winners, here the single best and the single worst thread, are taken into account[9].

Furthermore, as no narrative thread is ever *guaranteed* to succeed, it is always possible that no threads succeed and that the story produces no effect. For this reason, if there were to be no negative (positive) suspense values for any of the threads, then we would still always set the lowest (highest) suspense value at *0*. The global suspense would then just becomes equivalent to the highest positive (lowest negative) suspense value. We therefore include zero in our minimum and maximum calculations in the following definition of the global suspense:

Definition 31 *Global suspense value*

$$globalSuspense = \max_{\forall Z \in N}\{Suspense(Z), 0\} - \min_{\forall Z \in N}\{Suspense(Z), 0\} \quad (4.38)$$

Using our scale of -10 to $+10$ for the Importance values of threads, we can see that the maximum suspense in our system would be $(+10) - (-10) = 20$. This would correspond to a case where both the very worst thing possible and the very best thing possible could happen to a protagonist with the highest degree of imminence.

[9]One potential argument for using this simple measure could be based on an idea of limited attention: we can only give our attention to the single most exciting positive thing and the single most exciting negative thing at any one time. The validation of this conjecture is left to future work.

This concludes the description of our mathematical model for calculating the suspense level of a story at each story step.

4.6 Toy world example: Arthur and Brian

To illustrate concretely how our model works, we now examine a very simple Toy world example, made of a few possible events and very limited interaction between threads. To keep things simple, we will also only consider the effect of Imminence and Importance on the suspense levels in this story.

Suppose we have Arthur (A) who needs to escape from a tunnel which is n metres long. The last step brings him outside through a door. As soon as he is outside nothing more can happen to him, and his life is saved.[10]

However, we also have Brian (B) who is busily closing all the doors to the tunnels. He closes them one by one, and there are k doors left. He will reach the door to Arthur's tunnel in w doors time.

We can model this situation by creating the following two narrative threads:

```
A = [Arthur moves to (n-1)m away from door, Arthur moves to (n-2)m
    away from door,..., Arthur moves to 1m away from door, A gets
    out of the tunnel])
Importance(A)=10

B = [Brian shuts door 1, Brian shuts door 2, Brian shuts door 3,...,
     Brian shuts door w,..., Brian shuts door k ])
Importance(B)=0
```

We add the following disallowing event-pair to $\mathbb{D}$:

```
(Brian shuts door w, Arthur gets out)
```

[10] N.B.: Any similarity between this situation and the process of obtaining a Phd is entirely fortuitous.

This means: 'when the w^{th} door is shut, Arthur can no longer get out of the tunnel'.

We have put an Importance value on the storyworld state after the last event in thread A as +10, in other words, Importance$(A) = +10$. This means that in relation to his current situation, Arthur would have a much better situation once he is out of the tunnel and a free man. The Importance value of thread B can be set as 0, that is, it is trivial for the reader whether Brian succeeds in closing all the k doors. Now, globally, what is at stake in this story? If thread A succeeds then we have a positive value of +10, if thread B succeeds then we stay at **zero**. So in this story, there are 10 points at stake.

Now we can show how our model would deal with this situation as the story unfolds. We use the first option of a function for Imminence as follows:

Thread A:

- *Completion Imminence* of thread A is $\frac{1}{n}$, that is, it is inversely proportional to the distance n left to the door.
- *Interruption Imminence* of thread A is $\frac{1}{m}$, where m is the number of doors still left to close before Brian reaches the door to Arthur's tunnel. This means that the Imminence of the interruption or failure of thread A depends on events in thread B.

Thread B:

- *Completion Imminence* of thread B is $\frac{1}{k}$, where k is the number of doors left to close,
- *Interruption Imminence* of thread B is 0, that is, thread B cannot fail. In this storyworld, all the doors will be shut eventually.

To keep our example simple, we leave out the Foregroundedness and Confidence factors. The suspense S created by threads A and B will thus be simply equal to the Importance $\times$ Total Imminence:

$$S(A) = \text{Importance}(A) \times (0.7 \times (\text{Completion Imminence}) + 0.3 \times (\text{Interruption Imminence}))$$
$$S(B) = 0 \tag{4.39}$$

So we have the following:

$$S(A) = (+10) \times (0.7 \times \frac{1}{n} + 0.3 \times \frac{1}{m}) \tag{4.40}$$

and, using our working model of global suspense, this is also the *Total suspense* created by the story after each event. Now we can set the total number of doors as 5, and $w = 4$, that is, the 4^{th} door closes Arthur's tunnel. Furthermore, Arthur's tunnel is $5m$ long. We can now tell the following story which we show together with the values for n, m and *suspense* in Table 4.1.

Table 4.1: Suspense values for the Toy world example

event	n	m	suspense
start	5	4	2.15
Arthur moves to 4m away from door	4	4	2.50
B closes door 1	4	3	2.75
B closes door 2	4	2	3.25
Arthur moves to 3m away from door	3	2	3.83
B closes door 3	3	1	5.33
Arthur moves to 2m away from door	2	1	6.50
Arthur moves to 1m away from door	1	1	10.00
A gets out of the tunnel	0	1	0

Let us now imagine some different configurations to explore the intuitions behind our model. Suppose that Arthur has **100m** to crawl out of his tunnel, and Brian has **10 doors** to close:

$$\text{Suspense} = 10 \times (0.7 \times \frac{1}{100} + 0.3 \times \frac{1}{10}) = 0.37 \tag{4.41}$$

Suppose now that Brian has **100 doors** to close and Arthur has only **4m** to go. The suspense will vary as follows as Arthur crawls towards the door: **1.78, 2.36, 3.53, 7.03**.

For the case where Arthur has **100m** to crawl, and Brian **100 doors** to close, the suspense will be **0.1**.

So the Imminence and therefore the perceived suspense grows when either outcome approaches, and grows the most when both threads are nearing their individual final events or their point of interaction.

Chapter 5

Applying the model to a domain

We now describe the steps we took to create a computational implementation of our model which we could apply to a particular storyworld to generate suspense predictions.

5.1 A computational implementation of our suspense model

Our suspense model follows a three-step process.

Firstly, a new story event is told and the storyworld interpretation of the effect of this event is computed. In our model, this results in changes in the states of narrative threads.

Secondly, we compute the potential contribution of each narrative thread to the suspense felt at that moment in the story. This individual contribution is either positive for narrative threads that are seen as producing desired effects in the storyworld, or negative for those seen as producing undesired effects.

Thirdly, once we have all these values, we concentrate on the most positive and most negative suspense values ignoring all others. We determine the spread of these two values and take this as a measure of the global suspense level at that point in the story.

5.1.1 Implementation structure

We developed a PROLOG program to implement our mathematical model. The program has the following overall structure:

Each time a new event is told in the story:

A Update the set of narrative threads according to the effect of new event.

B Calculate the suspense contribution for each active narrative thread.

C The suspense level at this point in the story is then the difference between the maximum and minimum suspense values of all active threads.

We now show a simplified pseudo-code of our implementation. A fully detailed pseudo-code showing the corresponding PROLOG functions next to each step, is shown in Appendix A.1.1. The full program, storyworld and story files can be downloaded from http://www.richarddoust.eu/thesis/.

Simplified pseudo-code

First, acquire the new story in the form of a ordered list of events. Then, each time a new event α is added to the story, do steps A, B and C:

A DO THREAD MAINTENANCE DUE TO NEW EVENT

1. ADJUST FOREGROUNDEDNESS FOR ALL THREADS (*attentioncycle*): Lower the foreground value of all threads. In the next steps, threads linked to the new event will get their Foregroundedness value reset to the maximum level (1).
2. MATCH AND SHIFT EVENTS IN THREADS (*matchandshift*): For threads which contain the new event α, shift events into their Conveyed list up to α
3. EQUALISE CONVEYED EVENTS IN THREADS (*equalise*): Ensure that any newly conveyed events get placed into the conveyed list of all threads which contain them and give all threads which contain α the maximum foreground value
4. DEACTIVATE DISALLOWED THREADS (*disallowarcs*): Deactivate all threads that are disallowed by α
5. ACTIVATE NEW THREADS (*newarccheck*): Activate any *new* threads that contain α
6. CALCULATE THREAD CONFIDENCE LEVELS (*toldconflicts*): Update the confidence level for all threads based on the number of conflicted implicated prior events and the number of told events in the thread using the conflicting-to-told ratio ϕ.
7. CREATE NEW ACTIVE UNCONFIRMED THREADS (*newpredictedarcs*): Find and activate threads which contain events that

are in the Unconveyed events of confirmed active threads, even though these events have not yet been told in the story

8. DEACTIVATE COMPLETED THREADS (*completedarcs*): Deactivate threads which the new event α completes

B CALCULATE INDIVIDUAL SUSPENSE CONTRIBUTIONS

For each active thread Z (*calcsuspensefromarc*):

1. Find the **Number of steps to completion** (*findstepstocompletion*):
2. Number of steps to completion:
   If the thread Z is *confirmed*, use the number of its Unconveyed events as the **Number of steps to completion**
   - Otherwise, if the thread Z is *unconfirmed* (it has no told events), then we must first find the *Number of steps to (its) confirmation* and add this to the *Number of steps to completion*:

     Number of steps to confirmation of Z:

     (a) Find the non-empty set of active *confirmed* threads Y that could confirm Z, and use the lowest possible number of steps to the confirmation of Z from these as the **Number of steps to confirmation of Z**.

     (b) If there is no *confirmed* thread that could confirm Z, find the non-empty set of active *unconfirmed* threads Y' that could confirm it, find the minimum of all the *Number of steps to confirmation* for all Y' and add this number to the lowest possible number of steps to the confirmation of Z as above to get the *Number of steps to confirmation of Z.*

3. Find the **Number of steps to interruption** (*findstepstointerrupt*):

   - As in 4.29, we first define $R(Z, Y)$, the Interruption Imminence Number R of thread Z by any active thread Y:

$$
\begin{aligned}
\text{For all threads Z,Y, where } \text{state}_n(Z) = &\ (\text{active}|\text{inactive}, C_z, U_z),\\
\text{state}_n(Y) = &\ (\text{active}, C_y, U_y),\\
\text{if } (u_y, u_z) \in \mathbb{D}, u_y \in U_y &\text{ and } u_z \in U_z,\\
\text{then, } R(Z, Y) = &(\text{number of elements in } U_y \text{ before } u_y) + 1\\
\text{otherwise } R(Z, Y) = &\infty
\end{aligned}
\tag{5.1}
$$

   We then use the minimum number of events to interrupt Z as the **Number of steps to interruption**:

$$
\textbf{Number of steps to interruption} = \min_{\forall Y \in N} (R(Z, Y)) \tag{5.2}
$$

   - Otherwise, find the set of *unconfirmed* threads Y with at least one unconveyed predicted event which could disallow some predicted event in Z,
   Find all other threads X which could confirm each member Y of the set,
   For all X and Y, calculate the number of steps for X to confirm Y and for Y to interrupt Z
   Use the minimum value of the *sums* of these two numbers as

the **Number of steps to interruption of Z**

4. **Calculate suspense level for thread** Z (*suspensealgorithm*):

- Calculate the Completion imminence based on the **Number of steps to completion**
- Calculate the Interruption imminence based on the **Number of steps to interruption**
- Calculate the Total Imminence[1] (*imminencefunction*):

$$\begin{aligned}\text{Total Imminence} = {} & \rho.(\text{Completion imminence}) \\ & + (1-\rho).(\text{Interruption imminence})\end{aligned} \tag{5.3}$$

- Calculate suspense contribution of thread Z[2]:

$$\begin{aligned}\text{Suspense} = {} & (\text{Total Imminence}). \\ & (\text{Importance Value}). \\ & (\text{Foregroundedness}). \\ & (\text{Confidence})\end{aligned} \tag{5.4}$$

C CALCULATE GLOBAL SUSPENSE LEVEL AT THIS STORY STEP

From list of all suspense values for all threads, calculate global suspense G for this point in the story (*calcglobalsuspense*):

$$\text{globalSuspense} = \max_{\forall Z \in N}\{\text{Suspense}(Z), 0\} - \min_{\forall Z \in N}\{\text{Suspense}(Z), 0\} \tag{5.5}$$

[1] In our implementation, $\rho = 0.7$, boosting the effect of *Completion imminence*. Note that if ρ were set to 0.5, then the relative effect of Completion Imminence and Interruption Imminence would be the same.

[2] Note that this could be a negative number.

REPEAT Find the next event to tell in the story and repeat.

5.2 Writing a test-story

To test our implementation, we designed and wrote a short suspenseful story where an important judge drives towards his home with a bomb ticking in his car. This story, henceforth called the *Mafia story*, was inspired by the story used in Brewer and Lichtenstein's experiment (Brewer and Lichtenstein, 1982). To create step-by-step suspense level predictions for the story, the story needed to be split up into story steps or events. We followed Zwaan's protocol (Zwaan et al., 1995) mentioned previously, by splitting the story each time there had been a significant change in either time, space, interaction, subject, cause or goal. Here are the first few sentences of the story:

> Gianni was tired and dreaded the 15-minute drive home
>
> Taking on the Mafia in court was a tough, exhausting job
>
> He got into his old Lamborghini as the Town Hall clock struck six
>
> Just across the street a man in sunglasses was watching Gianni's car
>
> He pulled a remote control device out of his pocket and pressed a button on it
>
> The remote control screen started to flash: 10:00, 9:59, 9:58 ...
>
> A soft ticking noise started up at the back of Gianni's car
>
> Gianni drove out of the carpark ...

The full version of the story can be found in Appendix B.1.

5.3 Modelling the Mafia storyworld

The next step was to create the storyworld information which would enable our implementation to generate predictions about suspense levels for this story. Our model requires the following three types of information:

- $\mathbb{N}$, the set of narrative threads[3].
- Importance(Z), the importance values for all threads Z. In our system these values ranged from -10 to $+10$.
- $\mathbb{D}$, the set of pairs of mutually disallowing events

As we discussed previously, our model is designed to rely on information in a form which could be generated automatically from real-world data or corpora. The actual generation of this information lay however outside the scope of this research. We therefore created the narrative threads $\mathbb{N}$, their importance values Importance(Z) and the set of disallowing events $\mathbb{D}$, by hand, partly modelling our work on the event chains described in Chambers and Jurafsky (2009).

5.3.1 Constructing the narrative threads

To create the narrative threads, two questions need to be answered:

- which events in the storyworld should be linked together in a causal and/or intentional chain?
- with which events should a given narrative thread begin and end?

[3] We used the completeness rule that we mentioned in the previous chapter (see Rule 4 on page 111). This is the additional constraint that all events occur in at least one narrative thread. This means that $\mathbb{E}$, the set of possible events in the storyworld, contains at the very least the set of all events in all narrative threads.

Linking events in a thread

To guide the creation of the narrative threads for our storyworld, we reviewed our Mafia story for the presence of the following phenomena:

Habitual or stereotypical behaviour We tried to detect when characters were following habitual or stereotypical sequences of events.

Strong emotions If an event triggered a strong emotion for one or more of the characters, we took this as an indication that an explanatory narrative thread was needed.

Changes in knowledge If an event triggered a big change in the state of knowledge for a given character, we also took this as an indication that a narrative thread was needed.

We thus built up sequences of events based on our own expectations of what might happen in the storyworld at each step. We used as a guide the conditions on causal consequence inferences to verify that the events we linked together in a narrative thread had at least one of the following:

- a strongly supporting context,
- a strongly directive context.

Determining beginnings and end-points

To determine what events should *begin* a narrative thread, we used the following two criteria:

- The presence of a degree of surprise,
- A lack of natural or typical preceding events.

Candidates for the first event δ of a narrative thread had to trigger a certain degree of surprise. If they did not, and we could find other events that could precede them, either causally or intentionally, then we included these preceding events in a (longer) thread *containing* δ. We continued checking events in a thread for either a degree of surprise or a lack of typical preceding events. Events which satisfied both (or at least one) of these conditions could be taken as the starting event of a narrative thread.

Similarly, to determine what events should *end* a narrative thread, we used the following criteria:

- A significant change in the storyworld,
- A feeling of closure, or a lack of typical following events.

If an event in a narrative thread produced a significant change in the storyworld for one of the characters and there appeared to be no typical follow-up event, then we took this event to be the final event of a thread.

The criteria we have described for beginning and end-points are related to the conditions that the GLAIVE system uses for causal chains and intentional paths (see 3.2.5). We recall our summary of these conditions here:

- No event in a causal chain can negate the preconditions of another event in that chain
- A character must consent to all steps in a intentional path and intends the final effect of the last step during all the preceding steps

These conditions can both be seen as additional criteria for determining when a narrative thread should start or stop.

In our model, only the *final* events of a narrative thread are decisive in ascribing a value to the thread they complete. The choice of the final event is therefore a critical element in our storyworld modelling. It appears to us that, at least for the purposes of measuring suspense, in addition to the above, events which entail major changes in a character's state of knowledge should be taken as the final event of a narrative thread and therefore worthy of an importance value. Such events also often signal both the end of one thread and the beginning of another. They usually have important effects on the future events of a story and may even change the actual parameters within which the story evolves.

An example narrative thread

Guided by these criteria, we hand-crafted a series of narrative threads to define a storyworld in which the Mafia story could take place. Here is an example of one of the narrative threads we produced:

> Someone wants to kill Gianni →
> they plant a bomb in his car →
> they check that Gianni gets in the car →
> they trigger a remote control device →
> the countdown of the bomb starts on the remote control →
> the countdown starts in the car too →
> the countdown goes on for some time →
> the countdown reaches the end →
> the bomb explodes →
> Gianni gets killed

As we can see, this narrative thread includes elements of an **intentional path** (the intention of the person triggering the bomb), and parts of a **causal chain** (the countdown and explosion of the bomb).

Encoding narrative steps and events

It quickly became apparent that our natural language story would have to be encoded in a way that our theoretical narrative model had not predicted. Despite our efforts to split the story into steps according to Zwaan's protocol (Zwaan et al., 1995), it became clear that some of the sentences chosen produced several causal effects and could therefore affect several narrative threads at the same time. Here is an example from our story a few story steps from the end:

> He walked into his house and shut the door.

This story step indicates that Gianni has arrived home, changing the state of the 'going home' thread. However, in relation to our story, it also indicates that he has now left his car, and that he is now far away from his car (where the bomb is hidden), thus also changing the state of the 'exploding bomb' thread. In terms of its *effect on the story*, the actual event could be paraphrased in the following way:

> He left his car, he moved away from the bomb, he walked into his house and he shut the door.

Apart from appearing very unnatural in most story-telling environments, such a sentence would have to be split into three or four story steps using Zwaan's protocol.

To go any further, we needed to adapt our theory and create a distinction between the concept of a narrative *step* and a narrative *event*. A narrative step (of which the above sentence is an example) may contain one or more narrative events. Narrative steps are what makes up the *actual narration*: in a written story, these are usually the individual sentences, in a film, these are the individual film shots which can also include several events happening at the same time. For example, the following 10 story events (chosen from some fabula):

$$\text{STORY} = [e_1, e_2, e_3, \ldots, e_{10}]$$

might actually get narrated by the following sequence of 6 narrative steps:

$$\text{NARRATION} = [\{e_1\}, \{e_2, e_3, e_4\}, \{e_5, e_6\}, \{e_7\}, \{e_8\}, \{e_9, e_{10}\}]$$

Here we can see that some events have been grouped together and we may wonder if their chronological order has been respected. In the case of a sentence, as we can see in our example above, there may be a degree of chronological ordering that is retained through the order of the words in the sentence. However, it may be that the actual order of events *within* each narrative step is not important to the future development of the story, and that this is actually a criteria for the construction of such a narrative step. We will leave, however, the further development of this distinction to future work (see also 7.2.1).

We thus created a simple coding technique to allow our implementation to account for such cases, which of course occur very frequently in natural language stories. We first encoded all the story steps by creating event labels (in the following example *e30* is the label for the 30^{th} event):

```
event(e30,'He walked into his house and shut the door').
```

We then used this event label to create a mapping function from the story step to a list of all the events that it includes:

```
mapping(e30,[enters_home, leaves_car, is_far_from_car]).
```

In this way, we could accommodate complex story steps such as this one and still have enough leeway to create a natural story-telling experience.

5.3.2 Fixing the importance values

A narrative thread that succeeds creates a new storyworld situation that can be evaluated. To ascribe an importance value to a narrative thread, we need to look at the effect of the thread after the occurrence of its last event in the thread on the state of the storyworld. As we discussed in our informal model, the first principle guiding this ascription is based on two factors:

- the positive (or negative) valence of the reader's level of sympathy (or antipathy) towards the main character involved in the event
- the degree of perceived desirability (or undesirability) of the state of the storyworld after the event from the point of view of that character

In the case of the 'bomb' narrative thread mentioned above (see 5.3.1), we assume that the reader has *sympathy* for Gianni (valence $= +1$), and that getting killed is *highly undesirable* for Gianni (importance $= -10$). For this example, therefore, the event *Gianni gets killed* is ascribed an importance value of $+1 \times -10 = -10$.

Two other factors guided our importance value ascriptions:

- Over and above their other effects on the story, we also assumed that a *large increase in knowledge* for a positively (or negatively) valenced character would be categorised as positive (or negative) by the reader.
- If there was no clear effect on the storyworld after the final event in a narrative thread, an importance value of *zero or close to zero* was ascribed, thus annulling or reducing this particular thread's potential to make a direct contribution to the global suspense.

5.3.3 Creating the mutually disallowing events

The last element to be created was the set of disallowing event-pairs. Such event-pairs can disallow *unidirectionally* or *bidirectionally*. Here is a unidirectional example:

$$\text{resolving the mechanical problem} \overset{disallows}{\Longrightarrow} \text{the car breaking down}$$

Here is a bidirectional example:

$$\text{getting away from the bomb} \overset{disallows}{\Longleftrightarrow} \text{getting killed}$$

Creating the disallowing pairs consisted of simply determining which events in our narrative threads were mutually incompatible, that is, both of them could not conceivably co-occur in the same story for this storyworld. The incompatibility could be based on causal and/or intentional reasons.

5.4 Computational representations

We now present the story and storyworld data structures used in our PROLOG program.

5.4.1 The story representation

A story is represented in the following form: getstory($[s_1, s_2, ..., s_m]$). Here is the beginning of our Mafia story as an example:

```
getstory([thinks_about_job, wants_to_go_home,
    checks_gianni_gets_in_car, triggers_remote_control,
    countdown_starts, countdown_starts_in_car, drives_home
    ,...]).
```

5.4.2 The narrative threads representation

In our PROLOG implementation, narrative threads have the following format: arcdata(Thread Label, Importance Value, [List of events]).

Our example narrative thread shown earlier (see 5.3.1) was thus represented together with its *importance value* (–10) in the following way:

```
arcdata(gianni_gets_killed,-10,
[wants_to_kill_gianni,plants_bomb_in_car,
checks_gianni_gets_in_car,triggers_remote_control,
countdown_starts,countdown_starts_in_car,
countdown_goes_on,countdown_goes_to_end,
bomb_explodes,gianni_gets_killed]).
```

Here are some more examples of the narrative threads we created[4]:

```
arcdata(enters_home1,2,
[wants_to_go_home,drives_home1,
drives_home2,drives_home3,
arrives_home,turns_off_motor,
gets_out_of_car,leaves_car, enters_home]).

arcdata(car_breaks_down,-2,
[mechanical_problem_with_car,
car_breaks_down]).

arcdata(resolves_mechanical_problem,2,
[part_of_car_was_loose,
car_goes_on_bumpy_road2,car_gets_shaken2,
 mechanical_problem_with_car,
strange_noise_from_car, hears_noise_from_car,
```

[4]The complete list is to be found in Appendix A.2.

```
wants_to_find_noise, stops_car1,
goes_towards_noise, sees_something,
sees_mechanical_problem,
resolves_mechanical_problem]).

arcdata(gets_away_from_bomb,6,
[bomb_in_car_gets_shaken,
bomb_in_car_changes_behaviour,
strange_noise_from_car,hears_noise_from_car,
wants_to_find_noise,stops_car1,
goes_towards_noise,sees_something,
sees_bomb,gets_away_from_bomb]).
```

5.4.3 The disallowing event-pairs representation

The disallowing event-pairs are represented thus: disallow(event$A$, event$B$). Here is an example:

```
disallow(gianni_resolves_mechanical_problem,
    car_breaks_down).
```

In our story, this means roughly: ‘if Gianni resolves the mechanical problem, then the car will not break down’.

We included the following PROLOG clauses to allow an easy way to represent *mutually* disallowing pairs:

```
disallow(A,B):- disallowtwo(A,B).
disallow(A,B):- disallowtwo(B,A).
```

Here is an example:

```
disallowtwo(gianni_gets_away_from_bomb,gianni_gets_killed)
    ~.
```

This encodes the following two meanings:

- 'If Gianni gets away from the bomb then he won't be killed'
- 'If Gianni gets killed, then he won't be able to get away from the bomb'

5.5 Adjusting the program parameters

Our implementation depends upon a number of fixed parameters which determine the *relative importance* of the effects of the variables in the model. Experimentation with a few very basic stories and also with the Mafia story led us to adopt the following values that we then used for all further experiments. The goal we followed in adjusting these parameters was to produce the maximum *variability* in the different effects that the program variables had on the suspense level at any point in the story. Each variable had to be able to have some effect on the global suspense level.

We describe here the name and value of each parameter, together with its PROLOG name for easy reference to the program.

5.5.1 Foregroundedness: the attention decay factor

This is the rate at which the level of **Foregroundedness** of a thread goes down when the thread is not explicitly evoked in a story step. The PROLOG variable, name and value we used were the following:

- Decay factor for Foregroundedness (*attentiondecay*): $\beta = 0.88$

This value appeared to us to be the sweet spot, producing a decay curve which allowed recently forgotten threads to still potentially influence the global suspense level, whilst rapidly reducing threads' influence on suspense if they were not evoked in some way over four or five story steps.

5.5.2 Imminence: the interruption-to-completion ratio

This number regulates the importance of the ‘future’ or upcoming story events of a thread. It determines the relative effects of **Completion Imminence** and **Interruption Imminence** on the total suspense contribution of a given thread. The value we used was the following:

- Interruption-to-completion ratio (*interruptcompletionratio*): $\rho = 0.7$

Equality between the two would be represented by $\rho = 0.5$; the slightly higher number we use boosts the importance of Completion Imminence with respect to Interruption Imminence.

5.5.3 Revelatory suspense: the conflicted-to-told ratio

This number regulates the effect of ‘past’ or *conveyed* story events of a thread. It determines the relative importance of *conflicted* prior events on the confidence level of a thread compared to the number of *told* events. The value we used was the following:

- Conflicting-to-told ratio (*conflictingandconfirmedratio*): $\phi = 1.5$

The number we used is greater than 1 which means that the conflicted prior events in a thread produce proportionally *more uncertainty* for that thread than its told events can alleviate, thus decreasing the Confidence level of the thread. This ratio is above all important in revelatory suspense story situations as it comes into play when there are conflicting *interpretations* about an event in the story.

5.5.4 Interruption imminence: the interruptibility limit

As we discussed in the presentation of our mathematical model, all threads have an **interruption imminence number** which varies during the telling of the story. A high interruption imminence number indicates that an event that could disallow the thread is 'many story steps away' from occurring in the story. The **interruptibility limit** fixes an upper bound on this interruption imminence number for all threads. It means that we cannot for example, have a thread that can only be interrupted in say 1000 story steps. The interruptibility limit thus represents a certain minimum level of uncertainty about *any* thread in *any* storyworld. In a perfectly modelled storyworld, of course, there should be no limit, that is, the limit would be infinite. This relatively low cut-off value simulates the *inherent incompleteness* of our narrative thread system compared to a reader's reactions to a story; the idea is that there could always be *some* unknown events not present in the storyworld model which could interrupt an active thread. The value we used was the following:

- Interruptibility limit (*highnumberofsteps*): $\tau = 7$

The value means that all threads are considered to be interruptible no later than 7 steps away from the present story step. The effect of this limit was to create a minimal low level of Interruption imminence for all threads at all times.

5.5.5 Summary of the program parameters

For ease of reference, we give here an overview of the possible ranges of the program parameters together with the value we have chosen:

- Foregroundedness decay factor: $0 < \beta \leq 1, \beta = 0.88$

- Interruption-to-completion ratio: $0 < \rho \leq 1, \rho = 0.7$
- Conflicting-and-told ratio: $0 < \phi \leq \infty, \phi = 1.5$
- Interruptibility limit: $0 < \tau \leq \infty, \tau = 7$

5.6 Acquiring domain knowledge

As with any work on a model domain, there was a need to acquire additional real-world knowledge about this particular Mafia storyworld in order to validate the data structures we had created. We specifically needed to *calibrate* the relative importance of the narrative threads and make other adjustments to the threads themselves. This section presents the experimental study we set up to acquire this specific storyworld knowledge.

5.6.1 Designing a study to calibrate our implementation

Choosing a measuring method

Instead of the *actual* suspense level felt by readers, which we assume could only be measured with the help of brain scanning devices, and possibly other physical measures such as pulse rate, muscular tension and so on, we decided to use **the reader's subjective appreciation of suspense levels**.

We set up an online interface to collect self-reported suspense levels which presented a story one sentence at a time to participants. After reading each story step, participants were asked to evaluate the suspense level of the story at that point. The interface then displayed the next sentence in the story.

The main issue in the design of this study was to find a way to obtain self-reported suspense levels that would interfere as little as possible with the reading of the story. The method used had to both enable very rapid input and be very intuitive in order to reduce to a minimum both the time taken and the cognitive load needed to carry out the suspense level evaluation. To produce valid results, we needed to find a method which would not unduly interrupt the narrative *flow* of the story.

An additional criterion was the need to allow readers to indicate sufficiently fine grain variations in their suspense evaluations.

Discrete scale reporting methods

The first self-reporting method we tried out used arrows pointing *up*, *down* and *horizontally* at the right side of the screen. After reading each new story step, participants had to click on one of these 3 arrows to indicate whether they thought that the suspense of the story had gone *up*, *down* or stayed the *same*. In a slightly refined approach, participants could respond to each new story step by pressing one of 5 keys in answer to the following question:

Has the suspense level:

- gone up a lot?
- gone up a little?
- stayed the same?
- gone down a little?
- gone down a lot?

On clicking, the screen then showed the next sentence in the story.

These methods were both quick and intuitive. However upon testing, the rather wide groupings of suspense fluctuations gave crude results which were also hard to compare between participants. An unexpected problem of positive bias also arose, as most of the time the participants judged the suspense to either go up or stay the same, and hardly ever to go down. This bias further reduced the variation in the results we obtained.

The magnitude estimation method

The method we finally adopted is based on the magnitude estimation method. This method appeared to allow the necessary rapidity, intuitive ease of use and fine-grained definition of the participants' reactions together with the possibility of a high degree of rigour in the treatment and analysis of the data.

Magnitude estimation is an experimental technique which asks participants to give numerical values over a **freely defined and changing continuous scale** to estimate the magnitudes of a given stimulus (see Stevens, 1975). It has been successfully used in many psychological and physical domains and also notably in judgements of linguistic acceptability (Bard et al., 1996, Cowart, 1997).

As the participants could not know before reading the story the level of suspense that might be reached, this method seemed pertinent because participants could always choose a higher number than their previously imagined maximum if they judged the suspense level to have gone even higher than their maximum level up to that point. Similarly, if the suspense suddenly dropped to a very low value, the participants could indicate this without any ambiguity by entering a zero or a value close to zero.

This is a major advantage of the magnitude estimation technique: it allows participants to modify the intuitive scale they are using even in the middle of the telling of the story. Thus even very unexpected and sudden changes in suspense levels could be accommodated.

The method was implemented online for the Mafia story by asking participants to provide, immediately after reading each story step, a freely chosen numerical value corresponding to their perception of the suspense level of the story at that point. They they pressed 'ENTER' and the next story step was shown on the screen. The full explanatory text presented to participants is shown in Appendix C, but here is an relevant extract:

- *After reading each sentence in a story, you will be asked to indicate whether you think or feel the suspense level has gone down, stayed roughly the same or gone up.*
- *You can do this by typing **a number of your choice** which indicates the suspense level that you feel **at that point in the story**. It is important not to judge individual sentences as suspenseful or not, but rather the state of the story at that particular moment.*
- *You can enter any number greater than or equal to zero to do this. There is no maximum value you can give. Zero means no suspense at all.*
- *The idea is not to think too long before giving a value. Try and stay concentrated on the story itself during the experiment.*
- *Once you have typed in a number, you press **ENTER** to move on to the next sentence.*

Pilot tests on our Mafia story showed that this continuous free-scale magnitude estimation method produced more fine-grained measurements of suspense levels than the previous discrete methods based on fixed choices. The results were also more amenable to statistical analysis because they were based on a continuous scale. Statistical analysis of a given participant's results could be based on the implied numerical scale that their inputs created. In this way all the participants' results could be easily compared.

A warm-up story

To ensure that the participants were both seriously engaged in the study and at least somewhat practised in the chosen self-reporting method, we used a warm-up story. Both the warm-up story and the Mafia story were created specially for the experiment.

The warm-up story tells the tale of a potential mugger getting ready to attack a man walking in a park and is presented in Appendix B.3.

All participants were first asked to run through the warm-up story before going on to read the Mafia story. In this way, we aimed to eliminate participants who produced spurious outlier results for the warm-up story, and exclude their results from our analysis.

This type of safeguard is no doubt all the more necessary for online experiments where there is little or no control over the environment surrounding the participants nor over their level of concentration and language ability. Nevertheless, our results showed no cases of outliers based on the warm-up story. The only practical effect for this story was to give participants a degree of training in the self-reporting technique.

To introduce the main story, we also added a few sentences at the beginning to set the scene and introduce the main character, as follows:

> The scene: The offices in Palermo were starting to shut up for the evening. Gianni Ramazotti walked out of the Town Hall building and walked towards the carpark. He had arrived at the office several hours earlier than the rest of the staff. Gianni was tired and dreaded the 15-minute drive home.

5.6.2 The study method

Participants

A link[5] leading to the experimental set-up together with an invitation to take part in the study was sent out by email and Facebook messaging over a two-week period. A total of 40 people from various countries, all self-identified native or fluent speakers of English, took part in the study. We did not collect information on age or sex from the participants.

Materials

The online interface created for the experiment presented the warm-up story and then the Mafia story to the participants, recording their step by step self-reported suspense ratings. It used HTML and PHP both to store the data produced and to navigate between the different screens. The main design criteria of the interface was to reduce distraction from the story reading process as much as possible, whilst still providing a clear and user-friendly space for the participants' suspense ratings. The full introductory text shown to all participants and some screen shots of sample story steps are shown in Appendix C.

Procedure

Participants first read an introductory text which gave instructions about the experimental procedure. They then read through the warm-up story, rating each story step for a perceived suspense level. Once this story was completed, they went through the steps of the Mafia story in the same way.

[5] http://www.richarddoust.eu/trip2/indexB.php?lang=english

Treatment of results

The raw suspense ratings obtained in the experiment for each participant were converted to z-scores[6]. We then calculated the **mean z-score** and the **standard deviation** of all participants' z-scores for each story step.

5.6.3 The calibration process

First, we used our suspense algorithm on the storyworld model described above to produce suspense level predictions for all the steps in the Mafia story. For easy comparison, we then converted the predicted suspense values to z-scores, treating the results from our suspense model in the same way as those of the participants.

Once we had obtained both predicted and experimental values for suspense levels in the Mafia story, we could examine the degree of match and mismatch for different sections of the story.

Calibrating the importance values

By changing the importance values of some narrative threads, we modified their influence on the global suspense value produced by our algorithm at each story step. In this way, without changing the suspense modelling technique present in our model, we were able to vary the suspense curve generated by our theoretical model and achieve a better fit to the experimental values we had obtained. This process was done incrementally, by making small adjustments on one importance value to increase its influence on the suspense

[6] The z-score (or standardised score) is a dimensionless quantity obtained by subtracting the population mean from an individual raw score and then dividing the difference by the population standard deviation. The z-score of a raw score x can be written as follows:

$$z = \frac{x - \mu}{\sigma}$$

where μ is the *mean* and σ is the *standard deviation* of the population.

levels at a point of discrepancy and checking its overall effect on suspense for the whole story. Once we could not get any closer to the experimental curve in this way, the process was stopped.

This somewhat *ad hoc* procedure, based on varying one parameter of our storyworld modelling, that is, the relative importance of the narrative threads, also serves here as a reminder of the limited goal of this study: to show a *possible* path for the testing of a suspense theory in a given storyworld.

Modifying some narrative threads

In addition to the calibration of the importance values of the narrative threads, we made some modifications to the narrative threads themselves. We did this for two distinct reasons:

Missing threads We realised that there was a potential story outcome (and hence, in the terms of our model, a missing thread) that we had not included in our model which was having an effect on the participants suspense ratings.

Lengthening threads We needed to change the imminence of completion or interruption of a narrative thread by lengthening it.

Missing threads As an example of a missing thread, we can look at the final events of the story, where Gianni rushes out to see the result of a big bang. Here are the relevant events:

32. Just as he was hanging up his coat, he heard an incredibly loud bang
33. He rushed out of the house and surveyed the scene
34. Where his car had once been, there was nothing more than a chaotic heap of mangled blackened metal. . .

There is an ambiguity about the big bang; neither Gianni nor we as readers know exactly what has happened. Our first set of narrative threads failed to take into account this phenomenon. The results from the study however revealed that the suspense levels increased at events 31 and 32 before dropping again at event 33. This is actually a case of revelatory suspense. We therefore created the following new narrative threads to model this situation:

```
arcdata(sees_only_things_blown_up,-4,[bomb_explodes,
    things_get_blown_up,
hears_a_big_bang, wants_to_see_source_of_bang,
    goes_towards_bang, sees_only_things_blown_up]).

arcdata(sees_people_and_things_blown_up,-6,[bomb_explodes,
     people_get_blown_up,
hears_a_big_bang, wants_to_see_source_of_bang,
    goes_towards_bang, sees_people_and_things_blown_up]).

arcdata(sees_car_damaged_in_accident,-4,[
    car_damaged_in_accident,hears_a_big_bang,
    wants_to_see_source_of_bang,goes_towards_bang,
    sees_car_damaged_in_accident]).
```

Due to the presence of the event *hears_a_big_bang* in all three of these threads, all three of these narrative threads get activated when the big bang occurs at event 31. However, these three threads also contain conflicting implicated prior events. This situation reduces the threads' *Confidence* levels (see our Confidence definition 20 on page 122) and is akin to the presence of competition between the threads. The result for event 31 is an increase in suspense due to the activation of these three threads which is tempered by the competition between them. At event 32, as we get closer to the completion of all three threads, the imminence of the threads plays a role, and the suspense level increases. Finally, at event 33, the 'damaged car' thread succeeds and the other two get disallowed, bringing the suspense level back down.

Lengthening threads In addition to missing threads, we found that some threads produced imminence values that rose too high too soon. Because in our model, we use a fixed probability of transition between each event in a thread, the only way to reduce the imminence of a thread is to lengthen it. As an example of this, we show the first version of the following narrative thread:

```
arcdata(enters_home,2,[wants_to_go_home,drives_home,
    arrives_home,turns_off_motor,gets_out_of_car,
    leaves_car,enters_home]).
```

As we can see, the step *arrives_home* occurs straight after *drives_home*. Of course, in some stories, we might want to tell the *arrives_home* event straight after the *drives_home* event. This was not the case for our Mafia story, however. This thread's imminence of completion, that is, the moment when Gianni gets safely back into his house, was too high compared to the experimental results. The imminence of interruption of the 'bomb thread' due to this thread was also too high.

Another issue was that the story mentions the drive home several times. Leaving the thread in this form would have led us to repeat the event *drives_home* any time driving home was mentioned in the story.

There were therefore two reasons which led us to make a change in this narrative thread:

- To create a storyworld model which modelled the study results more closely,
- To bring the storyworld model closer to the actual events that occurred in the story; the drive home is the essential time-structuring element in this particular story.

To achieve these goals, we inserted additional *drives_home* events as follows:

```
arcdata(enters_home,2,[wants_to_go_home,drives_home1,
    drives_home2,drives_home3,
arrives_home,turns_off_motor,gets_out_of_car,leaves_car,
    enters_home]).
```

The final calibrated suspense values

We present the final calibrated suspense values together with the actual experimental suspense values for the Mafia story, both in z-score form, side-by-side in Figure 5.1 on page 177. The vertical brackets around each value represent the standard deviation of the z-scores for that story step[7].

Here is a short selection of the final calibrated Importance values:

- enters home = +2
- gianni gets killed = −10
- countdown fails = +1
- car breaks down = −2

[7] To better show the *transitions* from the presumed starting point of zero suspense for the experimental results in z-score form, we plot an additional story step preceding the others.

- resolves mechanical problem = +2

The complete list of Importance values are included in the presentation of each narrative thread in Appendix A.2.

As might be expected considering the small number of narrative threads we used to model this storyworld, even after this calibration process, discrepancies between predictions and results remained. One example is the following sentence:

1. Taking on the Mafia in court was a tough, exhausting job

Here the experimental results show a jump in the suspense value, whereas the model predicted no change. Our hypothesis for this difference is that the mention of the idea *Taking on the Mafia* was enough to activate a number of high importance threads that our storyworld modelling had not included. Although it would have been possible to create and include such threads, in order to limit the complexity of our storyworld modelling, and thus the total number of threads used, we decided to accept this discrepancy which seemed to only affect one event in the story[8]

Another discrepancy occurs at the following sentence:

11. The car started to shake as it clattered over them (the potholes)

Here the predicted suspense value shoots up considerably, whereas the experimental value increases by only a small amount. Upon examination of

[8]Alternatively, a word such a 'mafia' may trigger suspense *about the types of narrative gesture of the narrator*, that is, that the reader may believe that the narrator is more likely to create storyworld situations which are suspenseful. This is a possible interpretation for we might call the 'general feeling of suspense' that such a word produces. If so, then a complete model of suspense here would have to be extended with inferences about specific types of threads, e.g. the type of thread that involves the mafia, in addition to inferences about specific threads. This is left for future work.

the model, this difference appeared to be due to the high level of confidence in the storyworld thread that links the shaking of the car to the explosion of the bomb. Again, the fact that in our simplified model the shaking of the car event occurs in one thread and not several (which would create a degree of revelatory suspense and weaken its effect), appears to be a sufficient explanation for this discrepancy.

For this storyworld and this experimental set-up, then, most discrepancies can be resolved by the inclusion of additional or more refined storyworld information. Indeed, further studies could investigate whether the degree of agreement between our suspense model predictions and experimental results such as these would provide a useful *criterion* for determining whether a sufficient amount of storyworld information has been collected for a given story. These discrepancies notwithstanding, the model generates suspense predictions which are in good agreement with the results, especially if we compare the *direction of change*.

We now show and discuss how, for different story phases, the predicted values of our model depend more or less on two parameters of our model : **Foregroundedness** and **Confidence**.

A confidence-based narrative phase

The setting and the first events in the story activate the following 'getting home' thread which roughly models the events leading on from Gianni's intention to go home that evening:

```
arcdata(enters_home1,2,[wants_to_go_home,drives_home1,
    drives_home2,drives_home3,
arrives_home,turns_off_motor,gets_out_of_car,leaves_car,
    enters_home]).
```

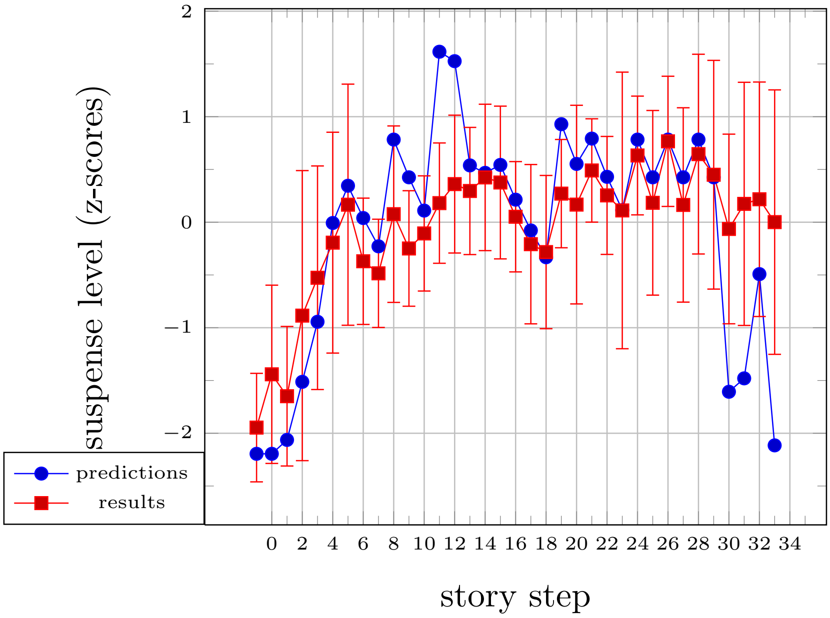


Figure 5.1: Calibrated predictions from our suspense model and experimental results for the Mafia-early story

A typical confidence-based phase occurs at the beginning of the story where we have a swift but progressive increase in the suspense value corresponding to the following story steps:

2. He got into his old Lamborghini as the Town Hall clock struck six
3. Just across the street a man in sunglasses was watching Gianni's car
4. He pulled a remote control device out of his pocket and pressed a button on it
5. The remote control screen started to flash: 10:00, 9:59, 9:58 ...
6. A soft ticking noise started up at the back of Gianni's car

Our model achieves part of this increase in suspense through the increase in completion and interruption imminence of the 'getting home' thread. However, the increase would not be great enough to match the study results, if it were not for the **revelatory suspense mechanism** which we have included in our model. In revelatory suspense, as we have seen, the confidence level of a given thread depends on *the number of implicated prior events it contains which are in conflict with other active threads.* In our storyworld model, as soon as the event *checks_gianni_gets_in_car* is told in the story, the following three threads get activated :

```
arcdata(gianni_gets_killed,-10,[wants_to_kill_gianni,
    plants_bomb_in_car,checks_gianni_gets_in_car,
    triggers_remote_control,countdown_starts,
    countdown_starts_in_car,countdown_goes_on,
    countdown_goes_to_end,bomb_explodes,gianni_gets_killed
    ]).

arcdata(gianni_tracked,-2,[wants_track_gianni,
    checks_gianni_gets_in_car, triggers_remote_control,
    map_on_remote_shows_giannis_position, gianni_tracked]).


arcdata(gianni_surprised,4,[wants_surprise_gianni ,
    checks_gianni_gets_in_car , calls_friends ,
    car_goes_to_surprise_point,friends_jump_out_on_gianni,
     gianni_surprised]).
```

We show a graphical plot graph of this situation in Figure 5.2 on page 180. However, the events preceding *checks_gianni_gets_in_car* in each of these three threads are in conflict:

```
wants_surprise_gianni
wants_track_gianni
wants_to_kill_gianni
```

The result is that at the start of the story, all three threads have implicated prior events which are in conflict with each other, and this reduces their possible suspense contribution (see the *Confidence* definition 20, on page 122). As the following story events occur, they disallow the conflicting narrative threads one by one, thereby increasing each time the suspense of the surviving threads, until only one remains: *gianni_gets_killed*.

A foregrounding-based narrative phase

If we examine the curves from events 16 to 19, both the predictions and the results show a steady decline. Here are the corresponding story events:

16. Suddenly, he saw something stuck in one of the wheelrims
17. He knelt down next to the wheel and carefully removed a rock that had got stuck there
18. Then, he got back in the car and drove off
19. On and on he drove over the pot-holed road

During this phase, no new threads are evoked and the only active conflicting threads which represent Gianni getting home: *enters_home* and Gianni getting killed: *gianni_gets_killed*, are not directly mentioned. Their contribution to suspense therefore goes down, decaying by the value $\beta = 0.88$ at each story step.

Decays in suspense occur from story steps 20 to 30 for similar reasons; each time the bomb is mentioned the *gianni_gets_killed* thread receives a

Figure 5.2: Revelatory suspense for the Mafia story: the strange man

K-links belong to the *wants_gianni_killed* thread,
T-links belong to the *wants_track_gianni* thread,
S-links belong to the *wants_suprise_gianni* thread.

Foregroundedness value of 1, the maximum. It is this mechanism that produces the series of suspense jags that we observe in this phase.

Discussion

A possible criticism of our experimental method could be the claim that *any* kind of suspense curve could be produced by manipulating the importance values given to the narrative threads.

Although it is true that the manipulation of the importance values can produce considerable differences in the predicted suspense levels, we claim a degree of *plausibility* for the values we have used. The potential death of Gianni, the main character, is given an importance of *–10*, his arriving safely home *+2*. These importance values are to be interpreted for the needs of narrative comprehension; they do not claim to correspond to the relative importances of such events in real life. However, there is a degree of correlation. There may be a way to translate real life importance values into their story-telling equivalents.

Another criticism could be that the idea of standardised importance values for narrative threads is impossible, participants' individual variations in these values will be too big to make the model valid; the death of the main character, say, may have more importance for one participant than for others and one might expect this difference to reveal itself in the suspense ratings they give at some of the story steps.

But in fact, our model of suspense shows a possible way to cater for this variation in suspense reaction between individuals: their differing relative evaluations for the storyworld events could be the result of their differing *suspense reactivity profiles.* One might be able to extrapolate individualised participant profiles by testing a whole series of stories with the same participant.

Nevertheless, the values we use to calibrate our storyworld modelling are based on the *average* reactions of all the participants. They are an attempt to model the most probable and most frequent suspense reactions for story situations in this storyworld: the most probable reactions from an average reader.

Another possible criticism is that the independent necessity of all the variables we use in our model has not been proven; there could be some variables whose effect on certain phases of the story is negligible, or whose modelling could be simplified.

However, the goal of this research was not to first discover a model of suspense and then prove the absolute necessity of all the model's variables. We had the more modest goal of showing a *possible* way to model suspense which does not structurally depend on specific storyworld information. The rigourous testing over a wide range of different stories of the necessity of all the model's variables is left for future work.

Finally, the process of calibrating the storyworld modelling to our Mafia story was instructive in itself. The study uncovered aspects of the storyworld which we had not been aware of using our own intuitive hand-crafted approach to creating the narrative threads.

Amongst others, we discovered the frequent need for revelatory suspense mechanisms which can be modelled by several threads sharing at least one event.

We also discovered the need for a definition of the relationship between what we have called *real-life event timing* and *narrative event timing* (see 5.6.3). It seems that the length of time that real-life events take does influence their timing in narrative. Similarly, there may be ways to translate real-life importance values into importance values useful for story-telling. Exactly how these relationships could be modelled is a subject for future research.

Study results such as these can have a creative feedback effect on storyworld modelling processes. We can easily imagine that after several such studies, more precise and structured methods for the construction of narrative thread storyworld models could be developed.

Chapter 6

Evaluating the model

6.1 Suspense predictions for the Mafia-late story

Central to our model of suspense is the possibility of predicting different suspense values for the different phases of a story. To evaluate the model, we needed a simple, constrained way to use it to predict suspense levels for a novel narrative situation. Of course, such predictions would be possible if we had a complete model of a new storyworld, but creating such a model is an extremely complex undertaking and lay outside the scope of our research. We decided to concentrate on testing the different suspense levels corresponding to *variations in the order* of a fixed number of story events in our calibrated Mafia storyworld. We could thus test like with like and limit the number of independent variables.

To do this, we created a variant of our Mafia story which differs only in that the vital information suggesting the presence of a bomb in the judge's car is revealed at a later point in the story. Apart from this change in the order of the events, we strove to create a story which used exactly the same events. To maintain the plausibility of the new story variant, we had to make minor changes in some event details, and two specific events could no longer be included, but our goal was to not activate any new narrative threads. We will henceforth refer to the new story variant we created as the *Mafia-late* story as opposed to the variant used in the first study, which we will call the *Mafia-early* story.

The full version of the *Mafia-late* story can be found in Appendix B.2. Here are the first steps of the story leading right up to the appearance of the 'man in sun-glasses':

> Taking on the Mafia in court was a tough, exhausting job.
>
> He got into his old Lamborghini as the Town Hall clock struck six.
>
> Gianni drove out of the carpark.
>
> He decided to take a shortcut which he knew would get him home in 8 minutes.
>
> He started to drive over a road full of potholes.
>
> The car started to shake as it clattered over them.
>
> After driving for a while, Gianni heard a strange noise coming from the back of the car.
>
> Gianni stopped the car and got out.
>
> A little worried, he walked towards the carboot.
>
> Suddenly, he saw something stuck in one of the wheelrims.
>
> He knelt down next to the wheel and carefully removed a rock that had got stuck there.

> Then, he got back in the car and drove off.
>
> Just across the street a man in sunglasses was watching Gianni's car ...

Thanks to the calibrated values for the narrative threads used to model the Mafia story-world that we obtained from the first study, we were in a position to create new predictions for this *Mafia-late* story variant. Using our PROLOG implementation, we produced the following predictions for the Mafia-late story which, for easy comparison, we show together with the *Mafia-early* suspense predictions in Figure 6.1 on page 186[1].

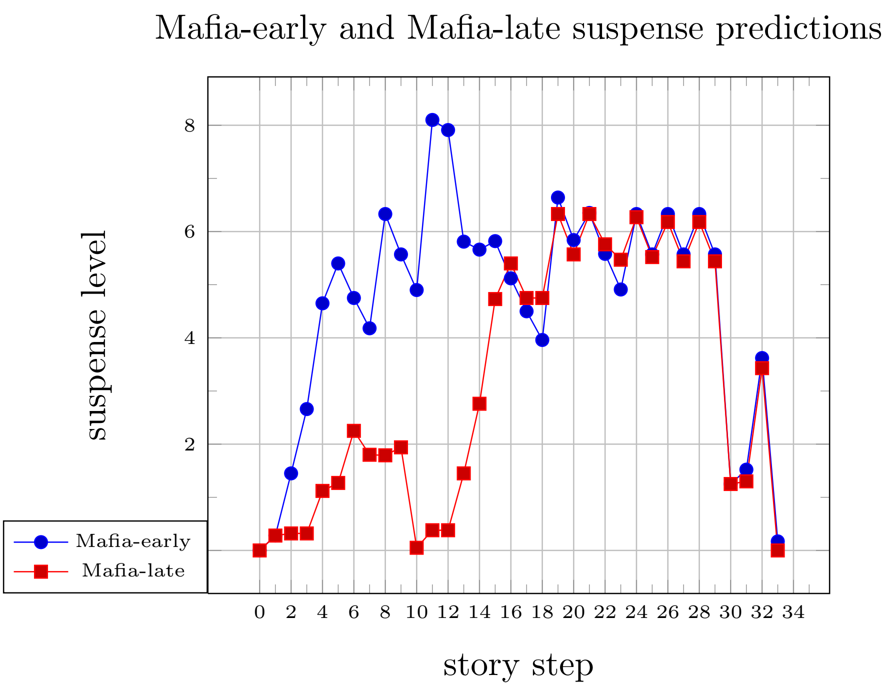


Figure 6.1: Predicted suspense values for Mafia-early and Mafia-late story variants

[1]The Mafia-late story has two events fewer than the Mafia-early story. To facilitate comparison, we have aligned the Mafia-late and Mafia-early results so that as far as possible the same events occur at the same point on the x-axis.

6.2 The suspense experiment

We were now in a position to test of our suspense theory using the predictions for the *Mafia-late* story variant based on the calibrated narrative thread values we had created for this storyworld. We conducted an experiment to obtain experimental suspense levels and compared these with the predicted levels.

6.2.1 Hypothesis

Our hypothesis was that the fluctuations in suspense levels predicted by our calibrated model for the Mafia-late story variant would agree with the fluctuations in the step-by-step averaged z-scores of story ratings for this story given by a new sample of participants. We used exactly the same experimental protocol and online interface as was used for the first calibration study.

6.2.2 Method

Participants

A link[2] leading to the experimental set-up was sent out by email and Facebook messages over a two-week period. A total of 46 people from various countries, all self-identified native or fluent speakers of English, took part in the experiment. We did not collect information on age or sex from the participants.

[2] http://www.richarddoust.eu/trip2/indexB.php?lang=english

Materials

The online interface was exactly the same as in our first calibration study. The introductory text for all participants and screen shots of sample story steps are shown in Appendix C and the full version of the story used can be found in the Appendix B.2. The participants' results were stored in an online file and analysed statistically using specially created PHP commands on the file.

Procedure

After clicking on the link to the experiment, each participant performed the following steps:

1. Reading an introductory text which gave instructions about the experimental procedure.
2. Reading through the warm-up story, rating each story step for a perceived suspense level.
3. Once this story was completed, the participant rated the 32 steps of the Mafia-late story variant in the same way.
4. The participants were invited to give some textual feedback on the experiment procedure and experience.

It was decided to stop the experiment once the number of complete results exceeded the number of participants in the calibration study (40). In fact, the results from 46 participants were found to be complete and were included.

The **independent** variables in this experiment were the Mafia-late story steps. The **dependent** variables were the suspense ratings given for each story step in the Mafia story. The control variable was the degree of English language fluency of the participants which was kept at a high level. The total time taken for the whole procedure was reported to be between 7 and 10 minutes.

6.2.3 Results and Statistical Analysis

The ratings obtained in the experiment for each participant were first converted to z-scores. For each story step, we then calculated the **mean z-score** and the **standard deviation of the z-scores** for all participants. For easy comparison, we also converted the *predicted* suspense values to z-scores. We present the comparison between the predicted and experimental results graphically in Figure 6.2 on page 190.

Observing the curves of the predictions and the results, we can see some good visual agreement between the two graphs. However, the vertical standard deviation values for the mean experimental z-scores at each story step are large, which also suggests relatively large fluctuations amongst participants. We now present some statistical analyses based on both the transitions between values and the absolute values to investigate the validity and reliability of these experimental results.

Pearson's and Spearman's correlations on the absolute values

Using the actual numerical values and not the transitions between values, we used predicted ratings (z-scores) for each **story step** and the **averaged z-scores** for the subject ratings. We show the **absolute values** used for the calculations of the Pearson correlations and the Spearman's correlations in Appendix D in Table D.1.

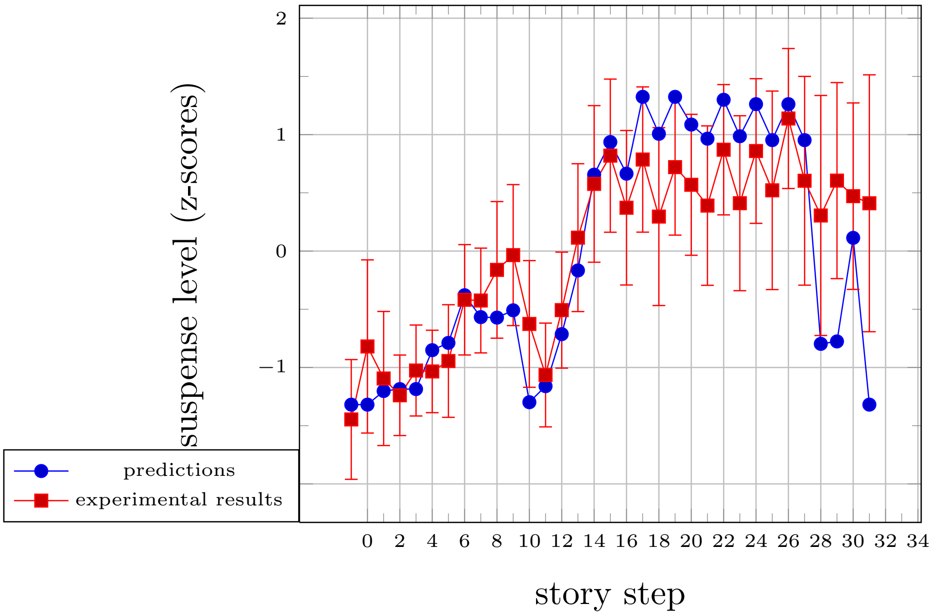


Figure 6.2: Experimental and predicted suspense for the Mafia-late story

The value of the **Pearson Correlation Coefficient** is **0.8234** and the **Spearman's Rho Coefficient** is **0.794**. Both values indicate a **strong positive correlation**.

Checking predicted transition categories

We also examined the ability of the model to correctly predict whether the suspense level *increased*, *stayed the same* or *decreased* at each new story step. We compared the direction of change of predicted and averaged experimental results for each story step. We found 24 correct predictions from a total of 32 transitions, or a prediction success rate compared to the averaged z-scores of the participants of **75%**.

We show the **absolute values** used for the calculations of the Pearson correlations and the Spearman's correlations in Appendix D in Table D.2.

Chi-squared and Fischer tests: integrating response bias

We noted considerable differences in the overall frequencies of the three different transition-types: *Up*, *Same* and *Down*. We determined the relative frequencies for each transition type which we show in the following Transition type frequencies table (Table 6.1).

Table 6.1: Transition type frequencies

Transition category	Totals	Percentage	Expected number of participants (N=46)
Ups	509	33.5%	15.42
Sames	710	46.8%	21.52
Downs	299	19.7%	9.06

This table shows that, if the suspense ratings were randomly distributed by category according to the overall frequencies, we should obtain roughly **15 Ups**, **22 Sames** and **9 Downs** at each story step. Only important departures from this distribution would indicate a tendency of subjects to favour or disfavour the same responses, or in other words, to agree. We used these average frequencies as the **expected distribution for a chi-squared test**.

To carry our this test, for all suspense level transitions in the story, we determined the *preferred responses* simply by determining whichever of *Up* or *Down* was the more frequent choice from the 46 participants for that transition, ignoring the overall response bias. We then took into account the overall response bias to test for the *significance* of the preferred responses using $p = 0.05$ and chi-squared > 3.84. We describe the method we used in detail in the Appendix Table D.3.

For the 32 story steps in this story, we found statistically significant results for 27 of them, which represents **84%**. It seems that there is a **high**

level of consensus in participants' ratings; this gives strong backing for the validity of the experimental set-up.

We then first performed a **Fischer's exact test** on the association between **predicted** and **observed** response categories. The association found is considered to be very statistically significant, showing *highly significant success in prediction.*

In another Fischer's exact test, we examined the association between the **correctness** of a transition prediction and its **significance** (or reliability). Here the test showed no significant correlation between the significance of a prediction and its correctness. In other words, for this experimental set-up, *reliable results were not more accurate than unreliable* results.

6.3 Discussion of experimental set-up

6.3.1 Experimental design

Our results suggest that we have found the right experimental form to connect people's perceptions of suspense in stories with the predictions generated by our model of suspense. We now review some of the choices we made in creating our model of suspense and the experiment we performed. Some of these choices can be seen as the limitations of this research and thus directions for future work.

Magnitude estimation

Standard uses for magnitude estimation methods are concerned with static evaluation of data, and try to elicit a sensitivity to ratios between different data points. This can sometimes occur through warm-up exercises where participants try to qualify the ratios between different pairs of straight lines for example. The participants then move on to the actual parameter under study and the hope is that this *feel* for the ratio between two different data events will be maintained during the experiment, at least by most participants most of the time.

Our use was less concerned with precise ratio ascription between different data points and more with the *flexibility* and *reactivity* that this method allows. We were trying to measure a parameter that is highly dynamic and depends enormously on context, in our case, on the story context made up of both the already told and upcoming story events in the mind of the reader. As the participants could not know what events were yet to occur in the story, they had to be allowed to *re-qualify* in relative terms the importance of certain events compared to others. Put simply, a difference in suspense between two events could be perceived as very important at one point in the telling of story, whereas later on, it might turn out, relatively speaking, to have been a very small difference, as new more important and suspenseful events are revealed.

This flexibility is one important quality that the magnitude estimation method provided. However, our use for this method went still further.

We believe that the method usefully enables comparison of potentially highly *idiosyncratic* and *diversified* scoring methods. This idiosyncrasy and diversity may also be an essential characteristic of perceptions of suspense by different individuals. In other words, it may well be that suspense is always a highly subjective reaction, which individuals perceive and qualify in different ways. By allowing a high degree of freedom in the scoring method, we hoped to mirror and *capture* the high degree of idiosyncrasy present in suspensefulness perceptions. Yet, the method also allowed the necessary statistical comparison and agglomeration of results.

Nevertheless, our suspense measurement process had a disadvantage: it was intrusive and surely created some interference with the reading process. Participants had to introspect about the suspense they felt, think of a number to represent this, type this number and press 'Enter' whilst at the same time attempting to stay immersed in the story they were reading. A better solution would not rely on such a conceptual task. We could perhaps use a physical lever to indicate variations in suspense in a more intuitive manner. Even then, there would be some interference. Interference could only be avoided by using some unobtrusive measuring system: a brain scan, or muscular tension and perspiration measurements. One would, of course, first have to show that there is a direct correlation between such measurements and perceived suspense, and this in itself is no small enterprise.

Narrative immersion

The raw material for our experimental procedure was the Mafia story, split up into sentence-sized chunks. The design criteria for this story were that it be short but nevertheless engaging enough to evoke suspense reactions from the readers which were as realistic as possible. To test our model, we were looking for a balance between a reasonable degree of both **computational tractability** and **narrative immersion**. Of course, with more resources it would be possible to use more complex stories. Our goal was however, not to write a brilliant story, but rather to test a theory in the most efficient way possible.

The feedback we collected from participants once they had finished carrying out the experiment showed that some had a degree of difficulty in really *engaging* with the story. Some of these participants mentioned ways in which they reacted to this difficulty:

- Some forced themselves to imagine the storyworld situations as vividly as possible,
- Some gave what they thought the response 'ought to be', and
- Some simply gave low suspense values.

Some participants also seemed to be using story understanding strategies which are not modelled in this experiment, predicting events in the story not from typical sequences of events but rather from memories of similar stories they had experienced. One participant described the falling away of suspense 'once it became obvious what was going to happen next in the story'. Yet other participants reported that they were more intrigued by the experiment than the story itself.

At least to some extent, all these reactions can be seen as a consequence of the limited material and time constraints available for this research. Clearly, it would be desirable to have richer, more involved, more complex, more unpredictable stories to test suspense on.

To create narrative immersion, many real-life stories often include long introductory chapters at the beginning which have the main goal of creating reader empathy for the characters of the story. Indeed, many parts of stories are concerned with creating and maintaining empathy. This was also the reason that our stories included a short paragraph setting the scene before the suspense experiment proper. The actual stories we created, however, had to be as short as possible in order to allow tractable modelling by a small number of narrative threads. In this light, this research can be seen as a starting point for future work which would aim for higher reader immersion levels by using longer more natural stories and more complex storyworld modelling.

Are we measuring suspense?

The question 'what exactly do you mean by suspense?' was, perhaps rather surprisingly, never asked in the feedback on the experiment by our particular cross-section of participants. Some participants' written feedback mentioned other concepts such as *danger*, *tension*, *fear* and *uncertainty* so we can suppose that these concepts may have also guided their responses. Because situations to which the concept of suspense is applied very often also evoke some of these other reactions, and also because most of us do not spend our time finely distinguishing between, say, the fear and the suspense that we feel whilst reading a story, all these concepts tend to be highly associated with one another.

It is also possible that *part* of some participants' reactions during the experiment may have been merely due to the *excitement* evoked by certain words. For example, perhaps a sentence with the word 'bomb' in it would produce a stronger reaction than the same sentence with the word 'explosive device' in its place. Of course, because the narrative threads used to model the storyworld would be the same in both cases, our model would predict the same level of suspense.

In our current experimental setup, however, all these effects were not visible, because we used the same words in the calibrating story as in the experimental story. More research on the cognitive and emotional meaning of certain isolated words would be needed to clarify this issue.

To conclude, one might expect that since individual *interpretations* (conscious or otherwise) of the concept of suspense might differ considerably, so too could the individual *suspense ratings* of events in a story. Considering the potentially very wide variation in the strategies used by the participants, due to possible cultural differences or idiosyncratic methods of self-reporting, the high statistical correlation between prediction and experiment is all the more striking.

To conclude, as with any experiment based on self-reporting, we cannot be absolutely certain that our set-up measures suspense. However, the high consistency of the results gives strong support that we are at the very least measuring *something very similar* for all participants.

6.3.2 Conclusions

Our model uses the concepts of **imminence**, **importance**, **foregroundedness** and **confidence** and proposes a way of combining these features to predict perceived suspense values. Our relatively simple experimental set-up based on self-reporting measurement of perceived suspense levels shows that our model of suspense is capable of successfully producing reliable predictions of suspense levels.

Perhaps surprisingly given the following three factors:

- the wide variation in participants' ratings,
- the relatively primitive experimental interface, and
- the short somewhat artificial nature of the stories used,

our experiment produced remarkably robust and consistent results. There are good fits between predicted values and participants' averaged scores and also between the predicted and experimental transition categories. We take these results as support for the validity of our model and for the suspense features it makes use of.

Chapter 7

Conclusions and future work

7.1 Our research question

We asked the following research question:

- What are the key components of a formal model of suspense that allows us to correctly measure and control suspense in narrative, whilst using a generic, domain-independent model of the story content?

We now examine the different aspects of this question and give our conclusions.

7.1.1 A formal model

In our review of the psychological, literary and computational literatures, we collected a set of concepts which we used to build up a model of suspense based on the concept of narrative threads. We constructed a mathematical formulation of our theoretical model and derived from that a computational implementation which we used to make suspense predictions for some simple stories.

As we have seen, our proposed way of structuring storyworld information using narrative threads, enables the modelling of future expected events, together with their degree of imminence and evaluations of their relative importance. It also includes a simple way to model variations in foregroundedness between different narrative threads. It is at least partially based on psychological models of narrative comprehension and can claim a degree of psychological plausibility.

7.1.2 Domain independence

Narrative threads can be used to produce a model of suspense for a given storyworld which is independent of the methods used to model inferential processes in the storyworld. The narrative thread structure does *not* depend on causal knowledge about particular domains, nor on information about human planning or goals. Such information is only necessary to provide the *content* of the narrative threads. Our model proposes a structural constraint on storyworld information which enables precise suspense predictions and yet is computationally tractable.

7.1.3 Validation through experiment

We have developed what we consider to be a useful experimental method based on magnitude estimation for obtaining self-reported suspense values for short stories. Our method used a specially created online interface and measures people's individual suspense evaluations as they read a given story.

We used this measuring technique to calibrate some model variables for a given storyworld using a short story. We then used our model to make suspense predictions for a variant of the first story and to test these predictions in an experiment.

The experimental results were in agreement with the predicted suspense levels to a high degree of statistical significance, showing that our suspense model achieves a high degree of *observational adequacy*.

Our results also suggest that we have found a useful experimental form for giving feedback about what features formal theories of narrative need.

7.1.4 A formula for suspense

We make the following points:

- In answer to our assumption from 1.2.4, there really *does* appear to be such a thing as a suspense profile for a given story about which many readers will be in agreement.

- Our model predicts variations in suspense during the step by step telling of a story which correlate with this suspense profile.

Recently, researchers have created a formula which successfully predicted happiness in a simple controlled experiment (Rutledge et al., 2014). As in our experiment, self-reporting techniques were used, but in addition participants' brains were scanned using functional magnetic resonance imaging. A surprisingly consistent relationship between rewards, expectations and happiness over a wide range of very different participants was found.

In a similar way, we have designed an experimental set-up which elicits suspense ratings for simple stories from participants. The validity of our experimental set-up suggests that we too can give a general formula for suspense:

Definition 32 *General suspense formula for one narrative thread*

$$\text{Suspense} = \text{Importance} \times \text{Foregroundedness} \times \text{Confidence} \times \\ (0.7 \times \frac{1}{\text{Events to Completion}} + 0.3 \times \frac{1}{\text{Events to Interruption}}) \tag{7.1}$$

The suspense due to a sequence of predicted events thus increases with the *importance* of the final event in the sequence, to the extent with which the sequence is *foregrounded* in the reader's mind and to the degree with which the reader is *confident* of having the right interpretation of the events in the story. The suspense level also increases as the number of remaining predicted events leading to the *completion* or the *interruption* of the sequence decreases.

7.1.5 Our contribution

The contributions this research makes include:

- A **domain-independent model of suspense** based on calculating the predicted conflict between narrative threads which extends the work of Brewer and Lichtenstein (1982).

- A **method for creating a computational model of a storyworld** based on the suspense model which uses narrative threads and event disallowing, and integrates causal and intentional storyworld information from diverse sources.

- A **suspense algorithm** which can use the storyworld model to create the suspense profile of a given story from that storyworld, by using the intermediate parameters of imminence, importance, foregrounding and confidence.

Our model has been validated in a limited domain of application by the empirical studies we carried out. The combination of imminence, importance, foregroundedness and confidence, has proven sufficient to model self-reported suspense levels in simple stories.

More generally, our model provides a potential starting point for the development of a formal and computational model of suspense which could be integrated into different narrative generation systems.

To summarise, the claim that our model is making is the following:

- The variation in perceived suspense levels during the telling of a story depends on list-like sequences of causally and intentionally predicted events and their varying levels of importance, imminence of completion, imminence of interruption, confidence and foregrounding.

7.2 Future work

Our model of suspense included a variety of simplifications. We now discuss some possible extensions of the model as pointers to future work.

7.2.1 Story events

We tested our theoretical model in only one particular medium: written stories, shown step by step. In our implementation, we used an extension to our theoretical model which distinguished between *simple* events and more *complex* narrative steps which group together small numbers of simple events (see 5.3.1). In fact, many story events can be understood as a grouping of several more fundamental events.

Our model could be extended to predict suspense for films and mixed media such as comic strips. One of our pilot studies involved the reordering of sequences from a Hitchcock film and it would be entirely feasible to use our narrative thread model to conduct an experiment based on suspenseful film sequences, similar to the text-based experiment presented in this research. Indeed, for film-based stories there may be additional challenges in modelling the storyworld, where, perhaps more so than in text-based stories, many things can happen in one story step.

Our distinction between narrative steps and simple events may not be sufficient to model the complexity of filmic narrative. It may be necessary to allow a given shot in a film to activate multiple narrative threads in parallel, much as if several written sentences had been told all at once.

The *interpretation* of a story in terms of events in a storyworld is a direction for future research.

7.2.2 Narrative threads

Automatic generation of narrative threads

In this research, the narrative threads used to model the storyworld were constructed by hand. Future work could attempt to link our suspense research to work on event chains by Chambers and Jurafsky (2009) and Li et al. (2013), and use automatic sourcing of corpora to generate a stock of narrative threads that could be used by our model to produce suspense predictions in a given storyworld.

It remains to be seen whether the event sequences typically generated by automatically sourcing corpora are sufficiently rich or detailed enough to adequately model simple storyworlds. Future work could also explore techniques for matching storyworlds with appropriate corpora.

Constructing narrative threads from different information sources

In general, narrative thread construction for storyworld modelling combines a variety of sources of information. We now discuss three sources that our model did not cover: logical reasoning, planning, story-specific threads. In creating the test stories for our model, we expressly avoided the need for narrative threads to be constructed in these ways. Clearly though, in general, stories do require them and further research should examine how the process of narrative thread construction from such sources could be formalised.

Logical reasoning In the Mafia story we used for our experiment, the logical reasoning necessary to follow the story was purposely kept to a minimum, so as not to create additional differences in participants' reactions due to their different reasoning strategies. But some story situations will require much more logical reasoning on the part of the reader for the conflicting suspenseful situations to be discovered. For the moment, we postulate the existence of an *external reasoning module* that carries out the necessary logical reasoning for a given situation and makes its conclusions available for the maintenance of the narrative threads.

Planning One obvious way in which our model could be connected to a planning approach to narrative would be to create and regularly update narrative threads based on characters' goals. There are different types of goals, but all contain *implicit or explicit sequences of events.* Such sequences can be translated into multiple narrative threads.

Story-specific sources A sequence of events which is created in the reader's mind *during the telling of the story*, possibly through repetition would be an example of a story-specific narrative thread which must be derived from scratch. Although the actual sequence of events of such a thread might be arbitrary - 'Sarah locked her front door after brushing her teeth' - subsequent repetitions of such a sequence in the story can potentially create a new story-specific narrative thread: every time Sarah brushes her teeth, we expect her next to lock her front door. Such a narrative thread is a kind of *generalisation on-the-fly* of a series of story events. In fact, any sequence of arbitrary events that occurs in a story can be used to create such a thread, which then becomes available to make predictions.

To deal with such cases, we postulate the existence of a *sequence detection module* that tracks the story for sequences, and upon their discovery, generates new story-specific narrative threads. We now briefly show how such a **sequence detection module** could be used to understand the story and joke-telling technique: the 'rule of three' we mentioned earlier (see 7.2.2) in 'Plot' (Dibell, 1988).

The 'rule of three' The 'rule of three' can be summarised in the following way:

1. We have a sequence of events A_1 leading to a final event X_1
2. We have A_2, a near-repetition of A_1 which leads to X_2, an event similar to X_1
3. Lastly, we have A_3, another near-repetition of event A_1 but this time leading to Z, a surprising variation of X_1.

Our *sequence detection machine* would thus tentatively put forward the narrative thread A', some generalisation of the sequence of events A_1, as a potential part of the library of available narrative threads for the storyworld. The second telling of A', A_2 would *reinforce* the identification of this particular sequence of events as a useful thread in this storyworld.

Now, once the sequence A' starts to be told for the third time, A_3, the listeners have very clear and strong expectations about what events at each part of the sequence A_3. However, the fact that A' is being told for the *third* time excludes the possibility that the story-teller wishes merely to reinforce the sequence. The story-teller must have some other unknown reason which means we have a case of revelatory suspense.

Thus we can see that the 'rule of three' is a narrative structure which is simple to construct but which automatically creates revelatory suspense about the story-teller's intentions. It is no surprise therefore that it is used in many levels and contexts as part of the human narrative tool-kit.

The computational tractability of revelatory suspense

Our model of revelatory suspense encounters some difficulties when we try to model it computationally: how can we model a whole range of unknown outcomes? Our pragmatic solution to this question has been to assume the existence of a limited range of known narrative threads which share common events and must be disambiguated. In this way, we can at least *simulate* the disambiguation process that occurs as the final outcome is revealed.

It may be however, that such suspenseful situations are better modelled by other means. We can imagine a class of suspenseful situations in which what we call *suggestive suspense* is present. Such situations can be created by strange events, that is, surprising events for which we have no immediate explanation and which thus combine surprise, curiosity and suspense. Finding effective ways to model such events in a computational model of narrative is a fertile research direction.

7.2.3 Importance values, timing and narrative thread detection

As we have seen in the discussion of the application of our theory to a domain (see 5.6.3), further research could explore the following:

- The relationship between real-life event timing and story-telling event timing.
- The relationship between real-life importance values and story-telling importance values.
- Formal ways to detect the necessity of certain narrative threads for a storyworld.

7.2.4 Suspense evaluation

Deriving an overall suspense value

Another refinement to our suspense modelling would be to derive **one overall suspense value** for a whole story. Such a value could be used to choose the most suspenseful variant from a series of story variations of the same story. If we plot a graph of the suspense level for each story step, then one model of the overall suspense felt during the whole story could the *area* under the suspense graph. However, it may be that other effects are important to the perceived overall suspensefulness of a given story, for example, the number of sudden short increases or suspense jags.

The spread in suspense values

The model of suspense we are putting forward uses what could be described as a first-past-the-post method for obtaining the suspense level at a given moment in the story: to calculate the degree of suspense, we find the *most positively valued thread* and the *most negatively valued thread*. We then use the spread in these two values as the global suspense level. This method neglects the effect of all the middle values which may nevertheless vary considerably.

This choice was made based on the premise that understanding a story is a task for which only limited attention is available, and that only the most important outcomes would be capable of creating suspense. However, for more complex models, it may be more useful to use the *statistical* spread in the individual narrative thread suspense values.

In addition, further experimentation using suspenseful stories that produce only two conflicting outcomes that are both positive (or negative) would also help determine whether suspense from different threads is additive whatever its valency, or whether we feel less suspense when only positive (or negative) outcomes occur in a given story. The results of such experiments could lead us to modify the current global suspense formula based on the spread in values[1]. This is also a question for future work.

Variable necessity and redundance

From this experiment alone we cannot deduce that all the variables we have included in our model are necessary to produce the complexity of the suspense profiles of the stories we tested. It may be possible to reduce some of the complexity of our theoretical model and still obtain a fit to our experimental results. Similarly, a more refined experimental set-up may reveal fluctuations in suspense levels which necessitate even more variables than those we have used.

[1] See also the discussion in 4.5.2

One area of exploration would be to identify whether certain variables in our model are more pertinent for certain situations. Story phases such as, for example, the culminating moments of a chase, could perhaps be modelled by a more basic formula. There are of course also advantages to using the same formula in all story situations.

The testing of both the *sufficiency* and *necessity* of each of our model's variables over a wide range of stories and storyworlds is a direction for future work.

7.2.5 Domains of application

Our model of suspense could be used to unpack certain story-telling strategies into their functional parts and hence explain their effectiveness and appeal. Here are a few examples:

Scene-switching: the power of 'meanwhile'

Scene-switching can be defined as the alternation between narrative viewpoints which show different sequences of events that belong to the same story. It is a technique that is ubiquitous in suspenseful film sequences.

One possible explanation for its use is that scene-switching increases the length of time that the suspense generated by a particular narrative thread is present in the story. If we have two narrative threads A and B, and we show first A then switch to B, then as long as A does not get forgotten, its suspensefulness can continue to affect the reader or viewer, even as we are watching events in narrative thread B. This a subject for future research.

Plot synopses

One domain where a suspense theory such as ours could be applied is written plot synopses and film trailers. Both of these seem designed to awaken curiosity about outcomes, either by the way the plot description is set up or by the way short suspenseful film sequences suggest a possible outcome but then switch rapidly away to a different scene.

7.2.6 The narrative cycle

The inspiration for our model came from work on suspense, curiosity and surprise by Brewer and Lichtenstein (1982). A key part of future work would therefore be to extend the narrative thread model to the concepts of curiosity and surprise in order to complete the typology of story-building elements. The experimental validation of our suspense model predictions suggests that similar experiments based on curiosity or surprise, could produce useful results.

As we have described in 3.5, surprise is a special case of disambiguation where the successful thread turns out to be one of the least expected, or in our formulation, one of the least *confirmed*. Because situations where curiosity is evoked necessarily involve unconfirmed threads, such situations are also often producers of surprise. Indeed, in our view, the key moments of many narratives are situations which combine all three entertaining narrative effects, creating curiosity to both generate suspense and allow surprise.

We sketch here what we dub the 'narrative breathing cycle':

- a suspenseful situation is *resolved* (completely or partially) by a surprising event
- the surprising event sets up the next revelatory or conflict-based sus-

pense phase,

- this phase leads after a certain time to another surprising event which resolves the suspense and so on.

We give a visualisation of this narrative cycle using narrative threads in Figure 7.1 on page 211.

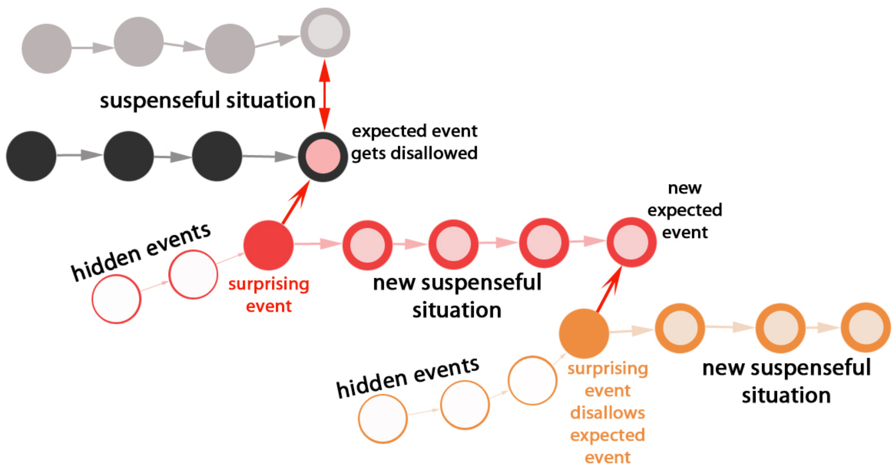


Figure 7.1: The narrative breathing cycle

Only the filled-in circles represent
events that are told in the story.
Each colour represents a different narrative thread.

Of course, stories differ in the amount of conflict-based suspense and revelatory suspense that they evoke and also in the length of time that suspense is maintained before a new surprising event occurs. We can say that stories have different suspense, curiosity and surprise profiles.

The precise modelling of the narrative cycle is a promising direction for future work.

7.3 Bridges to narrative-like domains

We believe that part of the power of our model lies in the structural constraint it imposes on information about storyworlds; we make the claim that our narrative thread model is sufficient to model suspense, wherever the information in these threads comes from. By abstracting the structure used to predict suspense from the content, we obtain a model that can be applied to many different fields, in fact, to any domain which is narrative-like, where there is a reader or listener or viewer who has expectations or predictions about an ongoing fluctuating process. The reader's predictions need to be based on some kind of grammar of event sequences which in our model we have called narrative threads.

We now review some potential domains of application of our model.

7.3.1 Suspense in music

Musical structure has often been compared to a narrative form (see for example Micznik, 2001) and the concept of suspense is also present in traditional musical theory in terms like a suspended cadence and a suspended fourth. Music can also evoke suspense over larger time scales as we feel the build-up of music towards a culminating high-point. When it finally arrives, the high-point also often contains an element of surprise, and this of course mirrors the behaviour of many story plots.

Future work could examine whether i) analogies from the musical world could give useful insights for our model of suspense in story, or ii) our model of suspense could be applied in a music theoretical setting. A description of music as narrative in the terms of our narrative thread theory raises several questions[2]:

[2]Similar questions could of course be asked of dance.

- What is a *musical* narrative thread and how can it be constructed? We assume that to model music, we need combinations of melodic, harmonic and rhythmic narrative threads.

- What would count as the beginning and end-points of a melodic, harmonic or rhythmic narrative thread and how could threads be evaluated?

- How are story-specific narrative threads introduced into the musical narrative by the exposition and the variations of a new theme for example. This question is analogous to our analysis of the 'rule of three' (see 7.2.2).

- What is the equivalent of a story character in musical narrative terms? An instrument? A *Leitmotif*?

Further research could also explore the role that *musical* suspense plays in creating and maintaining *story* suspense in films.

7.3.2 Analogies with linguistics

Earlier work on a net-linguistic implementation of an Earley Parser (see Schnelle and Doust, 1988, 1992) has had an indirect influence on the development of our model. There is a clear analogy between linguistic theories on sentence disambiguation and the model of suspense we propose. Narrative threads can be seen as analogous to lists of grammatical categories and the disambiguating process between narrative threads that occurs as a story is told can be seen as analogous to syntactic parsing. We can also draw a parallel between revelatory suspense and words that seems to belong to two linguistic categories. Surprise could be seen as analogous to the effect of garden-path sentences.

Story grammar systems have of course for a long time attempted to build on a parallel between linguistic and narrative structures. In contrast to the story grammar concept, however, our model does *not* use a fixed set of narrative categories. Instead, it uses causal and intentional list-like structures whose content comes from information about a specific storyworld. Suspense, curiosity and surprise are *meta-parameters* which can be extrapolated from the *intermediate parameters* of imminence, importance, foregrounding and confidence and these are in turn derived from the changing states of these list-like structures as a story unfolds.

A direction for future research would be to build a conceptual two-way bridge between linguistic theories and the narrative thread suspense model by treating sentences as a miniature stories made of 'word-events'[3]. For linguistics, some of the *intermediate parameters* could be the following:

- the measure of the varying likelihood of two different (or 'conflicting') words ending a given sentence as the sentence unfolds,
- the measure of the varying degree of ambiguity of interpretation of a word as the sentence unfolds.

7.3.3 Debugging the future: evolutionary benefits of suspense

A joke is a type of story that heavily uses suspense, surprise and curiosity. Recent work by Hurley, Matthew and Dennett (Hurley et al., 2011) grounds a new explanation of mirth and humour in evolutionary terms. We believe that an analogous development of suspense in narrative is possible.

Using the fact that we act boldly on committed beliefs, it became necessary, so goes Hurley et al.'s account, to catch potentially false beliefs before

[3]Hudson's Word grammar (Hudson, 2003) could be of interest in this light.

they became part of our long-term memory and we became committed to them. This led to the development of a kind of debugging mechanism which is applied to new incoming information. We can summarise their theory of humour in the following way:

- Humour is generated when information which has *covertly* been taken to be true, and to which we are epistemically committed, suddenly turns out to be false.

We can characterise surprise in a similar way:

- Surprise is generated when new information *overtly* enters a situation.

The main difference between surprise and humour lies in the fact that for surprise, we are not *covertly epistemically committed to a truth* which is then overturned by a surprising event.

We can characterise suspense in the same style:

- Suspense is generated when it is *overtly* recognised that only one of a series of predictions about a future situation can turn out to be true.

In Hurley et al.'s account we get more evolutionary rewards (laughing for example) for detecting *covert* errors than for *overt* ones. Nevertheless, as the enormous quantity of suspense stories testifies, we *do* get some debugging rewards for following and tracking suspenseful situations in stories. Future research could attempt to model suspense in terms directly linked to its evolutionary value.

7.3.4 A functional theory of narrative

This research raises the possibility of constructing a functional theory of narrative. Similarly to the claim made by Sternberg that we discussed in

2.3.3, such a theory would postulate that *all* the story steps that an author produces in telling a story must either modify, disallow or create at least one narrative thread and thus have some effect on the surprise, curiosity or suspense of the story at that point.

Thus, just as we can identify the roles that each word plays in a sentence, so we could obtain a precise description of the *role* that each story step plays in a given story. We could then derive a *functional narrative map* of a story which would be independent of the story's content, much like the syntactic analysis of a sentence.

7.4 Using concepts like suspense to guide narrative modelling

In some sense our work has been concerned with finding those constraints on storyworld modelling that enable predictions about suspense to be made. We would like to suggest that using suspense in this way could be a useful way to test the validity of theories of narrative. If then, through experiments, we obtain precise data about the suspensefulness of a given story, and the predictions of our narrative theory do not agree with this data, then we should be led to explore *either* where relevant storyworld information might be missing, *or* where our theory is lacking in some formal aspect. If suspense *is* one of the fundamental features of narrative, as we believe, then a good theory of narrative should make the calculation of the suspensefulness of a story easy; it should naturally fall out of the formalism used.

Scientific approaches to narrative changed dramatically upon the arrival of the computational era. The immediate challenge was to use computers to model everything, even something so intricate and multi-faceted as a story.

Perhaps we are now seeing a change in this approach and computational technology is now being used to attempt to focus more on the fundamental mechanisms which distinguish stories from other constructions. One of the differences between a simple sequence of events and a story is that the latter can create, amongst other effects, suspense. This research has attempted to tease out just how the creation of suspense can be modelled and explained from the sequence of events that is a story.

We see our work as a signpost towards further development of narrative models based on what we see as its fundamental ingredients. Further work could lead to explanations as to how higher-level narrative concepts, such as plot and character, naturally include some of the key ingredients of what is needed to build a successful story. We conclude that focussing on the basic components theoretically common to all stories is a fertile and necessary research path.

Appendix A

Computational implementation

A.1 Suspense algorithm

A.1.1 Detailed pseudo-code

First, acquire the new story in the form of a ordered list of events. Then, each time a new event α from the story is told, do the following steps:

A DO THREAD MAINTENANCE DUE TO NEW EVENT

1. FOREGROUNDING:
   - (*attentioncycle*)
     For all narrative threads $Z : \text{state}(Z) = (active, C_z, U_z)$, reduce the *Foregroundedness* of Z by the decay factor $\beta = 0.88$
   - (*reevoking*)
     For all narrative threads $Z : \text{state}(Z) = (active, C_z, U_z)$

and $\alpha \in C_z$,

set the *Foregroundedness* of Z to 1.

- (*activatesubthreads*)

  For all narrative threads Z : state(Z) = ($active, C_z, U_z$) ,

  (*activateembeddingthreads*)

  if there is a different active thread Y : state(Y) = ($active, C_y, U_y$),

  (*aretheylinked*)

  and there is a common member γ of C_x and C_y,

  or there is a common member γ of U_x and U_y,

  set the *Foregroundedness* of Y to 1.

2. MATCHING AND SHIFTING:

   - (*matchandshift*)

     For all narrative threads Z : state(Z) = ($active, C_z, U_z$),

     where $\alpha \in U_z$,

     shift all events up to and including α into list C_z

     and set the *Foregroundedness* of Z to 1.

3. EQUALISE: (ensure that any newly told events are in the told list of all threads)

   - (*equalisethreads*)

     For all narrative threads Z : state(Z) = ($active, C_z, U_z$),

     For all narrative threads Y : state(Y) = ($active|inactive, C_y, U_y$),

     if there exists $\gamma \in C_z$ and $\gamma \in U_y$

     then shift all events in U_y up to and including γ into list C_y

     and set the *Foregroundedness* of Y to 1.

4. DEACTIVATING THREADS

   - (*disallowarcs*)

for all narrative threads $Z : \text{state}(Z) = (\textit{active}, C_x, U_x)$,

and for each event $\gamma \in C_x$,

if γ has a mutual disallowing event λ,

set all threads Y : state$(\textit{active}, C_y, U_y)$ where $\lambda \in U_y$ to *inactive*.

5. ACTIVATING NEW THREADS:

   - (*newarccheck*)

     for all narrative threads $Z : \text{state}(Z) = (\textit{inactive}, C_z, U_z)$,

     if $\alpha \in U_z$,

     shift all events in U_z up to α into C_z,

     (*nopastconflicts*) if there are no past conflicts between C_z and other active threads not activated at this story step,

     (this condition is to allow multiple threads to be activated on the same event)

     and set Z to *active*.

6. CALCULATE THE CONFIDENCE LEVELS:

   - (*toldconflicts*)

     for all narrative threads $Z : \text{state}(Z) = (\textit{active}, C_x, U_x)$,

     set P as the number of events in C_x that have been told in the story,

     then for all narrative threads $Y : Y \neq Z, \text{state}(Y) = (\textit{active}, C_y, U_y)$,

     where $C_y \neq \emptyset$, that is, the thread has been *confirmed*,

     find the number of events Q in C_x that are in conflict with events in C_y

     $Confidence = \frac{1}{(1+\frac{\phi Q}{P})} = \frac{P}{(P+\phi Q)}$,

     where the Conflicting-to-told ratio $\phi = 1.5$

7. NEW ACTIVE UNCONFIRMED THREADS:

- $(newpredictedarcs)$

  for all narrative threads $Z : \text{state}(Z) = (active, C_z, U_z)$,

  for all narrative threads $Y : \text{state}(Y) = (inactive, C_y, U_y)$,

  if there are no active threads which conflict with U_y,

  then, if there exists an event $\gamma \in U_y : \gamma \in U_z$,

  then set $Y : \text{state}(Y) = (active, \emptyset, U_y)$,

  where Confidence level of Y is set to be the same as the Confidence level of Z.

8. DEACTIVATE COMPLETED THREADS:

- $(completedarcs)$

  for all narrative threads $Z : \text{state}(Z) = (active, C_x, \emptyset)$,

  set $\text{state}(Z) = (inactive, C_x, \emptyset)$

B CALCULATE INDIVIDUAL SUSPENSE CONTRIBUTIONS

For all narrative threads $Z : \text{state}(Z) = (active, C_z, U_z)$,

1. **Find the Number of steps to completion** for Z

   If thread Z is *confirmed*, set **Number of steps to completion** for $Z =$ the size of U_z.

2. Otherwise, if $\text{state}(Z) = (\text{active}, \emptyset, U_z)$,

   (that is, none of Z's events have been told yet),

   Find the non-empty set of active **confirmed** threads Y : $\text{state}(Y) = (\text{active}, C_y, U_y)$

   where $C_y \neq \emptyset$, and there exists some $\gamma \in U_y : \gamma \in U_z$,

   For all Y, find Before$_\gamma$, the number of events in U_y before γ,

   and After$_\gamma$, the number of events in Z after γ for Z to be completed,

   For all Y, **Number of steps to completion** $=$ Before$_\gamma$+After$_\gamma$.

Set the *Number of steps to completion of* Z as the *minimum* of all the above values over all Y.

Otherwise, find the non-empty set of active **unconfirmed** threads $Y' : \text{state}(Y') = (\text{active}, \emptyset, U_y)$,
where there exists some $\gamma \in U_y : \gamma \in U_z$
For all Y', find all $X : \text{state}(X) = (\text{active}, C_x, U_x)$ where there exists some $\zeta \in \mathrm{U}_x : \zeta \in U_y$ (and $X \neq Z$),
Find Before_ζ, the number of events still to go in U_x before ζ occurs,
Find Before_γ, the number of events in U_y between ζ and γ,
Find After_γ, the number of events in Z after γ occurs for Z to be completed,
For all X,Y', **Number of steps to completion** $= \text{Before}_\zeta + \text{Before}_\gamma + \text{After}_\gamma$.
Set the *Number of steps to completion of* Z as the *minimum* of all the above values over all Y' and X.

3. **Find the Number of steps to interruption** for Z (*findstepstointerruptfromonearc*)
For all *confirmed* threads $Y : Y \neq Z, \text{state}(Y) = (\text{active}, C_y, U_y) : U_y \neq \emptyset$
if there is an untold event $\epsilon \in U_y$ which disallows an untold event $\gamma \in U_z$,
(*stepstoevent*) Calculate the number of steps before Y can interrupt Z.
Then set the **Number of steps to interruption** as the *minimum* value of all the **numbers of steps to interrupt** for thread Z.

Otherwise, for all *unconfirmed* threads Y : state$(Y) = (\text{active}, \emptyset, U_y)$,

find an untold event $\epsilon \in U_y$ which disallows an untold event $\gamma \in U_z$,

(*findstepstoconfirminterruptingarc*)

find some $\mu \in U_x, \mu \in U_y$ for some **confirmed** active thread X,

where $X \neq Y, X \neq Z$,,

Then find number of steps in the thread to confirm the event:

Find the ranks $Rank_1$ of $\mu \in U_x$ and $Rank_2$ of $\epsilon \in U_x$

Set **number of steps in thread to confirm the event** to

$Rank2 - Rank1$ if $Rank_1 < Rank_2$, otherwise to 0.

Set **number of steps to interrupt** =

Steps to confirm the thread+Steps in thread to confirm the event.

Otherwise set **Number of steps to interrupt** to the maximum value (7)

Then set the **Number of steps to interruption of thread Z** as the *minimum* value from all the **Numbers of steps to interruption**.

4. **Calculate suspense level for thread** Z (*suspensealgorithm*):

   - Calculate the Total Imminence[1] (*imminencefunction*):

     Total Imminence = ρ.(Completion imminence)+$(1-\rho)$.(Interruption imminence)

   - Calculate suspense contribution of thread Z^2:

     Suspense = (Total Imminence).(Importance Value).

[1]In our implementation, the Interruption-to-completion ratio, $\rho = 0.7$, boosting the effect of *Completion imminence*. Note that if ρ is set to 0.5, then the relative effect of completion imminence and interruption imminence would be equal.

[2]Note that this could be a negative number.

(Foregroundedness).(Confidence)

C CALCULATE GLOBAL SUSPENSE LEVEL AT THIS STORY STEP

From list of all suspense values for all threads, calculate global suspense G for this point in the story (*calcglobalsuspense*):

$G = \text{Max}(\text{all individual suspense values}, 0) - \text{Min}(\text{all individual suspense values}, 0)$

REPEAT Then find the next event to tell in the story and repeat.

A.2 The Mafia storyworld data

A.2.1 The narrative threads

```
arcdata(enters_home1,2,[wants_to_go_home,drives_home1,drives_home2,
    drives_home3,arrives_home,turns_off_motor,gets_ouC_of_car,
    leaves_car,enters_home]).

arcdata(gianni_gets_killed1,-10,[wants_to_kill_gianni,
    plants_bomb_in_car,checks_gianni_gets_in_car,
    triggers_remote_control,countdown_starts,countdown_starts_in_car,
    countdown_goes_on,countdown_goes_to_end,bomb_explodes,
    gianni_gets_killed]).

arcdata(countdown_fails,1,
[countdown_starts_in_car,countdown_goes_on,countdown_fails]).

arcdata(things_geC_blown_up,-5,
[bomb_explodes,things_geC_blown_up]).

arcdata(people_geC_blown_up,-7,
```

```
[bomb_explodes,people_geC_blown_up]).

arcdata(gianni_tracked,-2,
[wants_track_gianni, checks_gianni_gets_in_car,
    triggers_remote_control, map_on_remote_shows_giannis_position,
    gianni_tracked]).

arcdata(gianni_surprised,4,
[wants_surprise_gianni, checks_gianni_gets_in_car, calls_friends,
    car_goes_to_surprise_point,friends_jump_ouC_on_gianni,
    gianni_surprised]).

arcdata(car_breaks_down,-2,
[mechanical_problem_with_car,car_breaks_down]).

arcdata(resolves_mechanical_problem,2,[parC_of_car_was_loose,
    car_goes_on_bumpy_road2, car_gets_shaken2,
    mechanical_problem_with_car, strange_noise_from_car,
    hears_noise_from_car, wants_to_find_noise, stops_car1,
    goes_towards_noise, sees_something, sees_mechanical_problem,
    resolves_mechanical_problem]).

arcdata(sees_only_things_blown_up,-4,
[bomb_explodes, things_geC_blown_up,hears_a_big_bang,
    wants_to_see_source_of_bang,goes_towards_bang,
    sees_only_things_blown_up]).

arcdata(sees_people_and_things_blown_up,-6,
[bomb_explodes, people_geC_blown_up,hears_a_big_bang,
    wants_to_see_source_of_bang,goes_towards_bang,
```

```
    sees_people_and_things_blown_up]).

arcdata(sees_car_damaged_in_accident,-4,
[car_damaged_in_accident,hears_a_big_bang,
    wants_to_see_source_of_bang,goes_towards_bang,
    sees_car_damaged_in_accident]).

arcdata(gianni_gets_killed3,-10,[wants_to_kill_gianni,
    plants_bomb_in_car,checks_gianni_gets_in_car,
    car_goes_on_bumpy_road1,car_gets_shaken1,bomb_in_car_gets_shaken,
    bomb_explodes,gianni_gets_killed]).

arcdata(gets_away_from_bomb,6,[bomb_in_car_gets_shaken,
    bomb_in_car_changes_behaviour,strange_noise_from_car,
    hears_noise_from_car,wants_to_find_noise,stops_car1,
    goes_towards_noise,sees_something,sees_bomb,gets_away_from_bomb
    ]).

arcdata(gianni_gets_killed3,-10,[plants_bomb_in_car,
    bomb_in_car_changes_behaviour,strange_noise_from_car,
    ignition_gets_triggered,bomb_explodes,gianni_gets_killed]).

arcdata(decides_and_enters_home,1,[wants_to_know_whaC_to_do,
    thinks_abouC_alternative1,thinks_abouC_alternative2,
    thinks_abouC_alternative3,decides_whaC_to_do,gets_ouC_of_car,
    leaves_car,enters_home]).

arcdata(car_damaged,1,
[car_crashes,car_damaged_in_accident,hears_a_big_bang]).
```

A.2.2 The disallowing event-pairs

```
disallow(A,B):- disallowtwo(A,B).
disallow(A,B):- disallowtwo(B,A).

disallow(wants_to_know_whaC_to_do, gets_ouC_of_car).
disallow(stops_car1, bomb_in_car_gets_shaken).
disallow(stops_car1, drives_home2).
disallow(turns_off_motor, bomb_in_car_gets_shaken).
disallow(resolves_mechanical_problem, mechanical_problem_with_car).
disallow(resolves_mechanical_problem, car_breaks_down).
disallow(sees_only_things_blown_up, sees_people_and_things_blown_up).

disallow(drives2, stops_car1).
disallowtwo(countdown_fails, bomb_explodes).
disallowtwo(sees_mechanical_problem, sees_bomb).
disallowtwo(gets_away_from_bomb, gianni_gets_killed).
disallowtwo(leaves_car, gianni_gets_killed).
disallowtwo(wants_track_gianni, wants_surprise_gianni).
disallowtwo(wants_to_kill_gianni, wants_surprise_gianni).
disallowtwo(wants_to_kill_gianni, wants_track_gianni).
disallowtwo(triggers_remote_control, calls_friends).
disallowtwo(countdown_starts, map_on_remote_shows_giannis_position).
disallowtwo(car_breaks_down, drives_home3).
disallowtwo(parC_of_car_was_loose, plants_bomb_in_car).
disallowtwo(mechanical_problem_with_car, bomb_in_car_gets_shaken).
disallowtwo(mechanical_problem_with_car,
    bomb_in_car_changes_behaviour).
disallowtwo(car_goes_on_bumpy_road2, bomb_in_car_gets_shaken).
disallowtwo(car_gets_shaken2, bomb_in_car_gets_shaken).
disallowtwo(bomb_explodes, car_damaged_in_accident).
```

```
disallowtwo(sees_only_things_blown_up, sees_car_damaged_in_accident).

disallowtwo(sees_people_and_things_blown_up,
    sees_car_damaged_in_accident).
```

A.3 The encoded Mafia stories

We give some contextual reminders after each event to facilitate comparison with the natural language story.

A.3.1 The Mafia-early story, used for the calibration study

```
getstory([thinks_abouC_job,
wants_to_go_home, % gets_into_car_aC_work % time 18:00
checks_gianni_gets_in_car,
triggers_remote_control,
countdown_starts, % countdown is 10 min
countdown_starts_in_car,
drives_home1, % arrives_home_in_8_min
countdown_goes_on, % countdown is 8:45
car_goes_on_bumpy_road,
car_gets_shaken, % strange_noise_from_car, hears_noise_from_car,
countdown_goes_on, % countdown is 2:59
wants_to_find_noise, % stops_car1,
goes_towards_noise,
sees_something, % sees_mechanical_problem,sees_rock,
    resolves_mechanical_problem,
drives_home2,
drives_home3,
countdown_goes_on, % countdown is 1:18
arrives_home,
```

```
countdown_goes_on, % countdown is 0:39
turns_off_motor,
wants_to_know_whaC_to_do,
countdown_goes_on, % countdown is 0:22
thinks_abouC_alternative1,
countdown_goes_on, % countdown is 0:13
thinks_abouC_alternative2,
countdown_goes_on, % countdown is 0:08
decides_whaC_to_do, gets_ouC_of_car,
leaves_car, enters_home,
hears_a_big_bang,
goes_towards_bang,
sees_only_things_blown_up]).
```

A.3.2 The story mappings to narrative thread events

```
mapping(e00,[thinks_abouC_job]).
mapping(e01,[wants_to_go_home]).
mapping(e02,[checks_gianni_gets_in_car]).
mapping(e03,[triggers_countdown]).
mapping(e04,[countdown_goes_on]).
mapping(e05,[bomb_remains_near_gianni]).
mapping(e06,[drives_home]).
mapping(e07,[drives_home]).
mapping(e08,[countdown_goes_on]).
mapping(e09,[car_goes_on_bumpy_road]).
mapping(e10,[car_gets_shaken]).
mapping(e11,[hears_noise]).
mapping(e12,[countdown_goes_on]).
mapping(e13,[wants_to_find_noise,stops_car]).
mapping(e14,[goes_towards_noise]).
mapping(e15,[sees_something]).
```

```
mapping(e16,[sees_rock]).
mapping(e17,[stops_looking,drives_home]).
mapping(e18,[drives_home,car_goes_on_bumpy_road]).
mapping(e19,[countdown_goes_on]).
mapping(e20,[arrives_home,stops_car]).
mapping(e21,[countdown_goes_on]).
mapping(e22,[wants_to_know_whaC_to_do]).
mapping(e23,[thinks_abouC_alternatives]).
mapping(e24,[countdown_goes_on]).
mapping(e25,[thinks_abouC_alternatives]).
mapping(e26,[countdown_goes_on]).
mapping(e27,[thinks_abouC_alternatives]).
mapping(e28,[countdown_goes_on]).
mapping(e29,[knows_whaC_to_do,gets_ouC_of_car]).
mapping(e30,[enters_home,leaves_car,is_far_from_car]).
mapping(e31,[hears_noise,bomb_is_far_from_gianni,bomb_explodes]).
mapping(e32,[goes_towards_noise,sees_something]).
mapping(e33,[sees_destroyed_car,discovers_noise_source,
    things_geC_damaged]).
```

A.3.3 The story mappings to natural language events

```
event(e00,'Taking on the Mafia in court was a tough, exhausting job
    ').
event(e01,'He got into his old Lamborghini as the Town Hall clock
    struck six').
event(e02,'Just across the street a strange man in sunglasses was
    watching Gianni s car').
event(e03,'He pulled a remote control device out of his pocket and
    pressed a button on it').
event(e04,'A countdown started on the screen of the device: 10:00,
    9:59, 9:58...').
```

```
event(e05,'At the same time, a soft ticking noise started up at the
    back of Gianni s car ').
event(e06,'Gianni drove out of the carpark').
event(e07,'He decided to take a shortcut which he knew would get him
     home in 8 minutes').
event(e08,'The countdown on the remote control device continued:
    8:45, 8:44, 8:43,...').
event(e09,'Meanwhile, Gianni started to drive over a road full of
    potholes').
event(e10,'The car started to shake as it clattered over them').
event(e11,'After a while, Gianni heard a strange noise coming from
    the back of the car').
event(e12,'The countdown on the device in the strange man s hand
    continued: 2:59, 2:58, 2:57 ...').
event(e13,'Gianni stopped the car and got out').
event(e14,'A little worried, he walked towards the carboot').
event(e15,'Suddenly, he saw something stuck in one of the wheelrims
    ').
event(e16,'He knelt down next to the wheel and carefully removed a
    rock that had got stuck there').
event(e17,'Then, he got back in the car and drove off').
event(e18,'On and on he drove over the pot-holed road').
event(e19,'The countdown on the device continued: 1:18, 1:17, 1:16
    ...').
event(e20,'At last Gianni turned into his street and pulled up in
    front of his house').
event(e21,'The countdown on the screen continued: 0:39, 0:38, 0:37
    ...').
event(e22,'He switched off the motor, leant back in his seat and
    sighed').
```

```
event(e23,'He wondered what he was going to do that evening').
event(e24,'The countdown on the screen continued: 0:22, 0:21, 0:20
    ...').
event(e25,'Perhaps he would get an early night').
event(e26,'The countdown on the screen continued: 0:13, 0:12, 0:11
    ...').
event(e27,'Or perhaps he should just order some pizza and watch TV
    ').
event(e28,'The countdown on the screen continued: 0:08, 0:07, 0:06
    ...').
event(e29,'Dreamily Gianni got out of the car and locked up').
event(e30,'He walked into his house and shut the door').
event(e31,'Just as he was hanging up his coat, he heard an
    incredibly loud bang').
event(e32,'He rushed out of the house and surveyed the scene').
event(e33,'Where his car had once been, there was nothing more than
     a chaotic heap of mangled blackened metal...').
```

A.3.4 The Mafia-late story, used in the online experiment

```
getstory([thinks_abouC_job,
wants_to_go_home, % gets_into_car_aC_work, time 18:00
drives_home1,
drives_home1, % arrives_home_in_8_min
% This event is missing from the Mafia-early version:
    countdown_goes_on, % countdown is 8:45
car_goes_on_bumpy_road2,
car_gets_shaken2, % hears_noise_from_car,
% This event is missing from the Mafia-early version:
    countdown_goes_on, % countdown is 2:59
wants_to_find_noise,% stops_car1,
goes_towards_noise,
```

```
sees_something, % sees_mechanical_problem,sees_rock,
    resolves_mechanical_problem,
drives_home2,
checks_gianni_gets_in_car,
triggers_remote_control,
countdown_starts, % countdown is 3 min
countdown_starts_in_car,
drives_home3,
countdown_goes_on, % countdown is 1:18
arrives_home,
countdown_goes_on, % countdown is 0:39
turns_off_motor,
wants_to_know_whaC_to_do,
countdown_goes_on, % countdown is 0:22
thinks_abouC_alternative1,
countdown_goes_on, % countdown is 0:13
thinks_abouC_alternative2,
countdown_goes_on, % countdown is 0:08
decides_whaC_to_do,
gets_ouC_of_car, leaves_car, enters_home,
hears_a_big_bang,
goes_towards_bang,
sees_only_things_blown_up]).
```

Appendix B

The stories in natural language

B.1 The Mafia-early story

<u>The scene</u>: The offices in Palermo were starting to shut up for the evening. Gianni Ramazotti walked out of the Town Hall building and walked towards the carpark. He had arrived at the office several hours earlier than the rest of the staff. Gianni was tired and dreaded the 15-minute drive home.

1. Taking on the Mafia in court was a tough, exhausting job
2. He got into his old Lamborghini as the Town Hall clock struck six
3. Just across the street a man in sunglasses was watching Gianni's car
4. He pulled a remote control device out of his pocket and pressed a button on it
5. The remote control screen started to flash: 10:00, 9:59, 9:58...

6. A soft ticking noise started up at the back of Gianni's car
7. Gianni drove out of the carpark
8. He decided to take a shortcut which he knew would get him home in 8 minutes
9. The remote control screen flashed: 8:45, 8:44, 8:43,...
10. He started to drive over a road full of potholes
11. The car started to shake as it clattered over them
12. After driving for a while, Gianni heard a strange noise coming from the back of the car
13. The remote control screen flashed: 2:59, 2:58, 2:57,...
14. Gianni stopped the car and got out
15. A little worried, he walked towards the carboot
16. Suddenly, he saw something stuck in one of the wheelrims
17. He knelt down next to the wheel and carefully removed a rock that had got stuck there
18. Then, he got back in the car and drove off
19. On and on he drove over the pot-holed road
20. The remote control screen flashed: 1:18, 1:17, 1:16,...
21. Eventually Gianni turned into his street and pulled up in front of his house
22. The remote control screen flashed: 0:39, 0:38, 0:37,...

23. He switched off the motor, leant back in his seat and sighed
24. He wondered what he was going to do that evening
25. The remote control screen flashed: 0:22, 0:21, 0:20,...
26. Perhaps he would get an early night
27. The remote control screen flashed: 0:13, 0:12, 0:11,...
28. Or perhaps he should just order some pizza and watch TV
29. The remote control screen flashed: 0:08, 0:07, 0:06,...
30. Dreamily he got out of the car and locked up
31. He walked into his house and shut the door
32. Just as he was hanging up his coat, he heard an incredibly loud bang
33. He rushed out of the house and surveyed the scene
34. Where his car had once been, there was nothing more than a chaotic heap of mangled blackened metal...

B.2 The Mafia-late story

The scene: The offices in Palermo were starting to shut up for the evening. Gianni Ramazotti walked out of the Town Hall building and walked towards the carpark. He had arrived at the office several hours earlier than the rest of the staff. Gianni was tired and dreaded the 15-minute drive home.

1. Taking on the Mafia in court was a tough, exhausting job
2. He got into his old Lamborghini as the Town Hall clock struck six

3. Gianni drove out of the carpark
4. He decided to take a shortcut which he knew would get him home in 8 minutes
5. He started to drive over a road full of potholes
6. The car started to shake as it clattered over them
7. After driving for a while, Gianni heard a strange noise coming from the back of the car
8. Gianni stopped the car and got out
9. A little worried, he walked towards the carboot
10. Suddenly, he saw something stuck in one of the wheelrims
11. He knelt down next to the wheel and carefully removed a rock that had got stuck there
12. Then, he got back in the car and drove off
13. Just across the street a man in sunglasses was watching Gianni's car
14. He pulled a remote control device out of his pocket and pressed a button on it
15. The remote control screen started to flash: 3:00, 2:59, 2:58...
16. A soft ticking noise started up at the back of Gianni's car
17. On and on Gianni drove over the pot-holed road
18. The remote control screen flashed: 1:18, 1:17, 1:16,...

19. Eventually Gianni turned into his street and pulled up in front of his house
20. The remote control screen flashed: 0:39, 0:38, 0:37,...
21. He switched off the motor, leant back in his seat and sighed
22. He wondered what he was going to do that evening
23. The remote control screen flashed: 0:22, 0:21, 0:20,...
24. Perhaps he would get an early night
25. The remote control screen flashed: 0:13, 0:12, 0:11,...
26. Or perhaps he should just order some pizza and watch TV
27. The remote control screen flashed: 0:08, 0:07, 0:06,...
28. Dreamily he got out of the car and locked up
29. He walked into his house and shut the door
30. Just as he was hanging up his coat, he heard an incredibly loud bang
31. He rushed out of the house and surveyed the scene
32. Where his car had once been, there was nothing more than a chaotic heap of mangled blackened metal...

B.3 The warm-up story

<u>The scene</u> : Jeffrey strolled out of his office and into the nearby park. He had been working hard, probably too hard, and he had a lot of things on his mind. He thought he would take a breath of fresh air and clear his head a little. He remembered the little café on the other side of the park.

1. Jeffrey decided to stroll down the winding path leading to the café.
2. In the park, Matthew had been wandering around the park for a while now and was getting nowhere.
3. Then he saw an office worker strolling along the main path through the park.
4. He looked a little absent-minded and Matthew could see the bulge of his wallet in the inside pocket.
5. Just then the man glanced over at him, and Matthew diverted his gaze, pretending to pick up a piece of rubbish.
6. When he looked up again, he saw the man walking down the path which led past a line of thick bushes in the middle of the park.
7. Matthew made his way towards the bushes.
8. Jeffrey continued ambling towards the café. It really was a beautiful day. He was looking forward to ordering a cappucino and just sitting for a while. He really needed to wind down. He should do this more often, he thought.
9. Matthew crouched down in the bushes next to the path. He could see the office worker approaching.
10. Jeffrey carried on down the path. He started to whistle. It was wonderful to get a breath of fresh air after a long morning in the office.
11. Matthew peered through the bushes. The office worker was whistling! This would be a piece of cake, he thought. He watched as he came up to the bushes.

12. Jeffrey walked from bush to bush. It was amazing how just being in the midst of nature made you feel better, he mused.
13. Matthew rolled up his sleeves and tensed his body. He was hardly breathing now.
14. Jeffrey came up to the last bush of the line, smiling and whistling.
15. Suddenly he stopped.
16. Where was his wallet ? Perhaps he had left next to his computer?
17. He would have to go all the way back to the office. It was so annoying.
18. But of course, there it was! He smiled and patted his breast pocket, relieved.
19. It was then that he heard a rustling noise.
20. He turned around and was just in time to see a figure jumping out at him from behind the bushes.
21. He threw up his arms to protect himself but the next thing he knew he was on his back pinned to the ground.
22. Matthew grabbed the office worker's wallet, but knocked his glasses to the ground in doing so.
23. Suddenly Matthew stopped, and looked more closely at the man.
24. A broad grin slowly spread across his face: "Jeffrey?", he chuckled.

Appendix C

Extracts from the online experiment

C.1 Introductory text

Suspenseful Stories: Testing a theory of suspense

To participate in this experiment, all you need to do is to read through two short stories sentence by sentence. The whole experiment should only take **5 minutes**.

After reading each sentence in a story, you will be asked to indicate whether you think or feel the suspense level has gone down, stayed roughly the same or gone up.

You can do this by typing **a number of your choice** which indicates the suspense level that you feel **at that point in the story**. It is important not to judge individual sentences as suspenseful or not, but rather the state of the story at that particular moment.

You can enter any number greater than or equal to zero to do this. There is no maximum value you can give. Zero means no suspense at all.

The idea is not to think too long before giving a value. Try and stay concentrated on the story itself during the experiment.

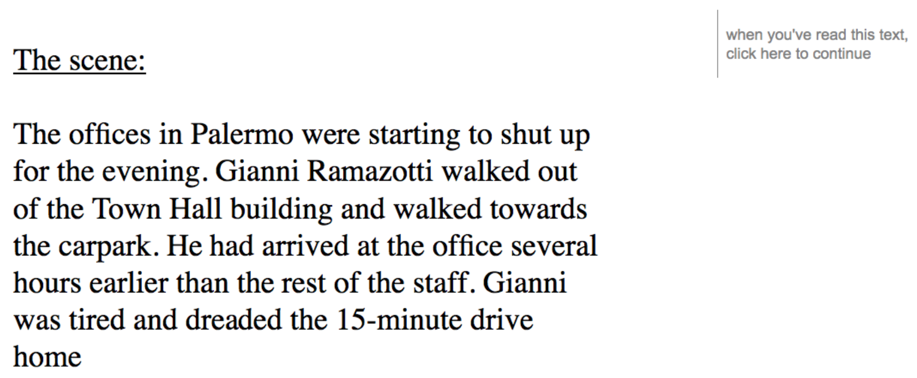
Once you have typed in a number, you press **ENTER** to move on to the next sentence.

After no more than about 15–20 sentences, when the story is finished, you will also be asked to give an **overall suspense level** for the whole story.

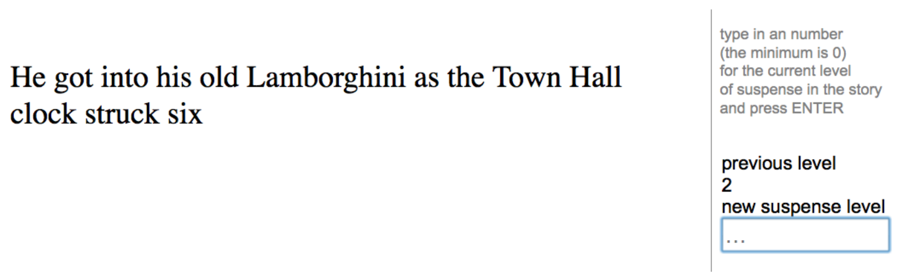
Then, you repeat the same procedure with a second and final story.

To get started, all you need to do is click on the big red **START->>** button!

C.2 Story setting screenshot



C.3 Story step screenshot


He got into his old Lamborghini as the Town Hall clock struck six
type in an number
(the minimum is 0)
for the current level
of suspense in the story
and press ENTER
previous level
2
new suspense level
...

Appendix D

Experimental results

D.1 Pearson and Spearman correlations tables

We show the **absolute values** used for the calculations of the Pearson correlations and the Spearman's correlations in Table D.1 on page 245.

D.2 Transition categories results

We show the success or failure of the transition categories in Table D.2. We found 25 correct predictions out of 31 transitions, or a prediction success rate compared to the averaged z-scores of the participants of **80%**.

Table D.1: Comparison of absolute values

predicted ratings	averaged subject ratings
0.00	–0.8
0.28	–1.09
0.32	–1.24
0.32	–1.03
1.12	–1.03
1.27	–0.94
2.25	–0.42
1.80	–0.42
1.79	–0.16
1.94	–0.03
0.05	–0.63
0.38	–1.06
1.45	–0.51
2.76	0.12
4.73	0.58
5.40	0.82
4.75	0.37
6.33	0.79
5.57	0.30
6.33	0.72
5.76	0.57
5.47	0.39
6.27	0.87
5.52	0.41
6.18	0.86
5.44	0.52
6.18	1.14
5.44	0.60
1.25	0.30
1.30	0.60
3.43	0.47
0.00	0.41

Table D.2: Table of transition categories

story step	prediction	direction	hit or miss
0	0.00		
1	0.28	UP	MISS
2	0.32	UP	MISS
3	0.32	SAME	MISS
4	1.12	UP	MISS
5	1.27	UP	HIT
6	2.25	UP	HIT
7	1.80	DOWN	HIT
8	1.79	DOWN	MISS
9	1.94	UP	HIT
10	0.05	DOWN	HIT
11	0.38	UP	MISS
12	1.45	UP	HIT
13	2.76	UP	HIT
14	4.73	UP	HIT
15	5.40	UP	HIT
16	4.75	DOWN	HIT
17	6.33	UP	HIT
18	5.57	DOWN	HIT
19	6.33	UP	HIT
20	5.76	DOWN	HIT
21	5.47	DOWN	HIT
22	6.27	UP	HIT
23	5.52	DOWN	HIT
24	6.18	UP	HIT
25	5.44	DOWN	HIT
26	6.18	UP	HIT
27	5.44	DOWN	HIT
28	1.25	DOWN	HIT
29	1.30	UP	HIT
30	3.43	UP	MISS
31	0.00	DOWN	HIT

D.3 Chi-squared method

1. For each transition we determined whether the 46 subjects overall preference was for *Up* or *Down* (ignoring *Same*). (If the number of *Ups* was the same as the number of *Downs*, we assumed a preference for the less frequent of the two, in this case *Down*.)

2. For each transition we computed chi-squared for the preferred category (either *Up* or *Down*) against the other two categories combined. We used expected values based on the overall frequencies for all story steps. Thus for *Up*, the expected value is **15.42** against **30.54** for *Down* and *Same* combined. For *Down*, the expected value is **9.06** against **36.94** for *Up* and *Same* combined. We then could calculate chi-squared for each story step based on the preferred response category. We show this in Table D.3. We then counted the total numbers of hits and misses and significant or non-significant results from this data.

3. Firstly, we performed an overall 2x2 Fischer's exact test for **association between predicted and observed response categories** (from the columns *Preferred response category* and *Predicted transition type* in Table D.3. We show the results of the Observed/Predicted Fischer test in Table D.4. For Fisher's exact test: the two-tailed P value equals **0.002**. The association between rows (groups) and columns (outcomes) is considered to be very statistically significant. This shows **highly significant success in prediction**.

4. Secondly, we compared the success for transitions in which the preference was *significant* (reliable) with the success where they were *insignificant* (from the columns *Hit* and *Significant* above). We show the results in Table D.5. For Fisher's exact test, the two-tailed P

value is **0.578**. The association between rows (groups) and columns (outcomes) is considered to be **not** statistically significant. This show that for this data there is no significant correlation between the *significance* of a prediction and its *correctness*. In other words, for this experimental set-up, the *reliable* results were not more accurate than the *unreliable* results.

Table D.3: Chi-squared table

story step	up	same	down	preferred response (no sames)	predicted transition	chi-squared for preferred response (p=0.05, chi-squared > 3.84)	significance	hit or miss
0	31	14	1	up	same	23.66	SIG	miss
1	4	25	17	down	up	8.66	SIG	miss
2	5	30	11	down	up	0.52	0	miss
3	17	26	3	up	same	10.91	SIG	miss
4	5	37	4	up	up	45.41	SIG	hit
5	8	37	1	up	up	45.41	SIG	hit
6	35	9	2	up	up	37.38	SIG	hit
7	10	32	4	up	down	26.80	SIG	miss
8	21	23	2	up	down	5.60	SIG	miss
9	15	25	6	up	up	8.94	SIG	hit
10	1	10	35	down	down	92.48	SIG	hit
11	0	18	28	down	up	49.30	SIG	miss
12	38	8	0	up	up	49.71	SIG	hit
13	38	8	0	up	up	49.71	SIG	hit
14	31	14	1	up	up	23.66	SIG	hit
15	22	23	1	up	up	5.60	SIG	hit
16	8	22	16	down	down	6.62	SIG	hit
17	24	21	1	up	up	7.17	SIG	hit
18	3	23	20	down	down	16.45	SIG	hit
19	26	17	3	up	up	10.91	SIG	hit
20	9	23	14	down	down	3.35	0	hit
21	7	25	14	down	down	3.35	0	hit
22	26	20	0	up	up	10.91	SIG	hit
23	5	21	20	down	down	16.45	SIG	hit
24	22	24	0	up	up	7.17	SIG	hit
25	6	25	15	down	down	4.85	SIG	hit
26	27	19	0	up	up	13.07	SIG	hit
27	4	18	24	down	down	30.67	SIG	hit
28	4	23	19	down	down	13.58	SIG	hit
29	17	18	11	up	up	0.65	0	hit
30	5	30	11	down	up	0.52	0	miss
31	10	21	15	down	down	4.85	SIG	hit

Table D.4: Observed/Predicted Fischer test

		observed			
		up	down	total	Fisher's exact test
predicted	up	14	4	18	Two-tailed P value
	down	2	10	12	= 0.00218522
	total	16	14		

Table D.5: Correctness/Reliability Fischer test

predictions		correct			
		yes	no	total	Fisher's exact test
reliable	yes	21	6	27	Two-tailed P value
	no	3	2	5	= 0.57786429
	total	24	8		

Bibliography

Abuhamdeh, S., Csikszentmihalyi, M., and Jalal, B. (2015). Enjoying the possibility of defeat: Outcome uncertainty, suspense, and intrinsic motivation. *Motivation and Emotion*, 39(1):1–10. [Cited on page 3.]

Alwitt, L. (2002). Maintaining attention to a narrative event. *Advances in Psychology Research*, 18:99–114. [Cited on page 34.]

Aristotle, t. S. H. B. (trans. 1974). *Poetics*, volume VII,X,XI. www.gutenberg.org. [Cited on page 17.]

Bae, B. and Young, R. (2008). A use of flashback and foreshadowing for surprise arousal in narrative using a plan-based approach. *Interactive Storytelling*, pages 156–167. [Cited on page 59.]

Bae, B. and Young, R. (2009). Suspense? surprise! or how to generate stories with surprise endings by exploiting the disparity of knowledge between a storys reader and its characters. *Interactive Storytelling*, pages 304–307. [Cited on page 59.]

Baggett, P. (1979). Structurally equivalent stories in movie and text and the effect of the medium on recall. *Journal of Verbal Learning and Verbal Behavior*, 18(3):333–356. [Cited on page 31.]

Bailey, P. (1999). Searching for storiness: Story-generation from a reader's perspective. In *Working notes of the Narrative Intelligence Symposium*, pages 157–164. [Cited on pages 40 and 55.]

Bal, M. (1997). *Narratology : introduction to the theory of narrative*. University of Toronto Press. [Cited on page 20.]

Barber, H. and Kudenko, D. (2008). Generation of dilemma-based interactive narratives with a changeable story goal. In *Proceedings of the 2nd international conference on INtelligent TEchnologies for interactive enterTAINment*, page 6. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering). [Cited on page 52.]

Bard, E. G., Robertson, D., and Sorace, A. (1996). Magnitude estimation of linguistic acceptability. *Language*, pages 32–68. [Cited on page 166.]

Baroni, R. (2007). *La tension narrative : Suspense, curiosité et surprise*. {P}oétique. Seuil. [Cited on page 37.]

Barros, L. M. and Musse, S. R. (2008). Towards consistency in interactive storytelling: Tension arcs and dead-ends. *Computers in Entertainment (CIE)*, 6(3):43. [Cited on page 55.]

Barthes, R. (1966). *Critique et vérité*. Siglo XXI. [Cited on page 7.]

Black, J. B. and Bower, G. H. (1980). Story understanding as problem-solving. *Poetics*, 9(1):223–250. [Cited on page 47.]

Boujarwah, F., Abowd, G., and Arriaga, R. (2012). Socially computed scripts to support social problem solving skills. In *Proceedings of the SIGCHI*

Conference on Human Factors in Computing Systems, pages 1987–1996. ACM. [Cited on page 65.]

Bower, G. H. and Morrow, D. G. (1990). Mental models in narrative comprehension. *Science*, 247(4938):44–48. [Cited on page 26.]

Boyd, B. (2005). Evolutionary theories ofart. *The Literary Animal. Evolution and the Nature of Narrative. Evanston*, 101:149–178. [Cited on page 3.]

Branigan, E. (1996). *Narrative comprehension and film*. Routledge. [Cited on page 20.]

Brewer, W. and Lichtenstein, E. (1982). Stories are to entertain: A structural-affect theory of stories* 1. *Journal of Pragmatics*, 6(5-6):473–486. [Cited on pages I, 7, 8, 12, 31, 33, 36, 48, 59, 60, 86, 150, 201, and 210.]

Bringsjord, S. and Ferrucci, D. A. (2000). Artificial intelligence and literary creativity: Inside the mind of brutus, a storytelling machine. *Computational Linguistics*, 26(4). [Cited on pages 47 and 48.]

Bruner, J. (1991). The narrative construction of reality. *Critical inquiry*, pages 1–21. [Cited on page 2.]

Cannon, W. B. (1932). The wisdom of the body. [Cited on page 4.]

Caplin, A. and Leahy, J. (2001). Psychological expected utility theory and anticipatory feelings. *Technology*, 55. [Cited on page 35.]

Carroll, N. (1984). Toward a theory of film suspense. *Persistence of Vision*, 1:65–89. [Cited on page 34.]

Carroll, N. et al. (1990). *The philosophy of horror: Or, paradoxes of the heart*. Routledge. [Cited on page 32.]

Carroll, N., Vorderer, P., Wulff, H., and Friedrichsen, M. (1996). Suspense: Conceptualizations, theoretical analyses, and empirical explorations. [Cited on page 34.]

Chambers, N. and Jurafsky, D. (2008). Unsupervised Learning of Narrative Schemas and their Participants. *Journalism*. [Cited on page I.]

Chambers, N. and Jurafsky, D. (2009). Unsupervised learning of narrative schemas and their participants. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 2-Volume 2*, pages 602–610. Association for Computational Linguistics. [Cited on pages 65, 104, 151, and 203.]

Chambers, N. and Jurafsky, D. (2010). A database of narrative schemas. In *LREC*. [Cited on page 65.]

Chatman, S. (1980). Story and discourse: Narrative structure in fiction and film. [Cited on page 49.]

Cheong, Y. and Young, R. (2006). A computational model of narrative generation for suspense. In *AAAI 2006 Computational Aesthetic Workshop, Boston, MA, USA*. [Cited on pages 58, 67, and 68.]

Cheong, Y.-G. and Young, R. M. (2008). Narrative generation for suspense: Modeling and evaluation. In *Interactive Storytelling*, pages 144–155. Springer. [Cited on pages 58 and 60.]

Collins (2003). Collins english dictionary, complete and unabridged. [Cited on page 6.]

Cowart, W. (1997). *Experimental syntax: Applying objective methods to sentence judgments*. Sage Publications. [Cited on page 166.]

Csikszentmihalyi, M. and Csikszentmihalyi, M. (1991). *Flow: The psychology of optimal experience*, volume 41. HarperPerennial New York. [Cited on page 2.]

de Wied, M. (1995). The role of temporal expectancies in the production of film suspense* 1. *Poetics*, 23(1-2):107–123. [Cited on page 4.]

de Wied, M., Tan, E., Frijda, N., Macar, F., Pouthas, V., and Friedman, W. (1992). Duration experience under conditions of suspense in films. *NATO ASI series. Time, action and cognition: Towards bridging the gap*, pages 325–336. [Cited on page 34.]

Dibell, A. (1988). Plot. elements of fiction writing. *Writer's Digest Books, Cincinnati, OH.* [Cited on pages 14 and 205.]

Fairhall, S. L., Albi, A., and Melcher, D. (2014). Temporal integration windows for naturalistic visual sequences. *PloS one*, 9(7):e102248. [Cited on page 75.]

Freytag, G. (1863). Die technik des dramas. *WA Shakespeare-Jahrbuch Bd*, (BAZ. 1863, in vier Folgen, 26. III.—9. IV):27–60. [Cited on page 17.]

Genette, G. (1972). Discours du récit, essai de méthode. *Figures*, 3. [Cited on pages 7 and 19.]

Gerrig, R. and Bernardo, A. (1994). Readers as problem-solvers in the experience of suspense. *Poetics*, 22(6):459–472. [Cited on pages 34, 36, 58, 59, 60, and 61.]

Goyal, A., Riloff, E., et al. (2013). A computational model for plot units. *Computational Intelligence*, 29(3):466–488. [Cited on page 50.]

Graesser, A. C., Singer, M., and Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological review*, 101(3):371. [Cited on pages 26, 28, and 29.]

Hasson, U., Landesman, O., Knappmeyer, B., Vallines, I., Rubin, N., and Heeger, D. J. (2008). Neurocinematics: The neuroscience of film. *Projections*, 2(1):1–26. [Cited on page 2.]

Heider, F. and Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, pages 243–259. [Cited on page 1.]

Hennequin, A. (1890). *The art of playwriting: being a practical treatise on the elements of dramatic construction; intended for the playwright, the student, and the dramatic critic*. Houghton, Mifflin. [Cited on pages C and 21.]

Hitchcock, A. (1956). *The man who knew too much*. [Cited on pages 20 and 25.]

Hodhod, R., Magerko, B., and Gawish, M. (2012). Pharaoh: Context-based structural retrieval of cognitive scripts. *International Journal of Information Retrieval Research (IJIRR)*, 2(3):58–71. [Cited on page 64.]

Hudson, R. (2003). Word grammar. [Cited on page 214.]

Hurley, M. M., Dennett, D. C., and Adams, R. B. (2011). *Inside jokes: Using humor to reverse-engineer the mind*. MIT press. [Cited on page 214.]

Jansen, A. S., Van Nguyen, X., Karpitskiy, V., Mettenleiter, T. C., and Loewy, A. D. (1995). Central command neurons of the sympathetic nervous system: basis of the fight-or-flight response. *Science*, 270(5236):644–646. [Cited on page 4.]

Jones, M., Love, B. C., and Maddox, W. T. (2006). Recency effects as a window to generalization: separating decisional and perceptual sequential effects in category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(2):316. [Cited on page 94.]

Kerby, A. P. (1991). *Narrative and the Self.* Indiana University Press. [Cited on page 2.]

Kintsch, W. (1980). Learning from text, levels of comprehension, or: Why anyone would read a story anyway. *Poetics*, 9(1):87–98. [Cited on page 31.]

Kintsch, W. (1988). The role of knowledge in discourse comprehension: a construction-integration model. *Psychological review*, 95(2):163. [Cited on page 22.]

Lämsä, A.-M. and Sintonen, T. (2006). A narrative approach for organizational learning in a diverse organisation. *Journal of Workplace Learning*, 18(2):106–120. [Cited on page 2.]

Lang, R. (1999). A declarative model for simple narratives. In *Proceedings of the AAAI fall symposium on narrative intelligence*, pages 134–141. [Cited on page 46.]

Lebowitz, M. (1985). Story-telling as planning and learning. *Poetics*, 14(6):483–502. [Cited on pages 51 and 85.]

Lehnert, W. G. (1981). Plot units and narrative summarization*. *Cognitive Science*, 5(4):293–331. [Cited on page 49.]

Lehnert, W. G. (1982). Plot units: A narrative summarization strategy. *Strategies for natural language processing*, pages 375–412. [Cited on page 49.]

Li, B., Lee-Urban, S., Johnston, G., and Riedl, M. (2013). Story generation with crowdsourced plot graphs. In *AAAI*. [Cited on pages 104 and 203.]

Liu, H. and Singh, P. (2002). Makebelieve: Using commonsense knowledge to generate stories. In *AAAI/IAAI*, pages 957–958. [Cited on page 65.]

Magliano, J., Kopp, K., McNerney, M. W., Radvansky, G. A., and Zacks, J. M. (2012). Aging and perceived event structure as a function of modality. *Aging, Neuropsychology, and Cognition*, 19(1-2):264–282. [Cited on page 24.]

Magliano, J. P., Loschky, L. C., Clinton, J. A., and Larson, A. M. (2013). Is reading the same as viewing? In B. Miller, L. C. and McCardle, P., editors, *Unraveling the Behavioral, Neurobiological and Genetic Components of Reading Comprehension*, pages pp. 78–90. MD: Brookes Publishing Co., Baltimore. [Cited on page 23.]

Mandler, J. M. and Johnson, N. S. (1980). On throwing out the baby with the bathwater: A reply to black and wilensky's evaluation of story grammars*. *Cognitive Science*, 4(3):305–312. [Cited on page 47.]

Mateas, M. and Stern, A. (2003). Integrating plot, character and natural language processing in the interactive drama façade. In *Proceedings of the 1st International Conference on Technologies for Interactive Digital Storytelling and Entertainment (TIDSE-03)*. [Cited on page 55.]

McNamara, D. S. and Magliano, J. (2009). Toward a comprehensive model of comprehension. *Psychology of learning and motivation*, 51:297–384. [Cited on page 26.]

Meehan, J. (1977). Tale-spin, an interactive program that writes stories. In

In Proceedings of the Fifth International Joint Conference on Artificial Intelligence. [Cited on pages 41, 43, and 67.]

Meehan, J. (1981). Tale-spin and micro tale-spin. *Inside computer understanding. Hillsdale, NJ: Erlbaum*. [Cited on pages 41 and 43.]

Micznik, V. (2001). Music and narrative revisited: Degrees of narrativity in beethoven and mahler. *Journal of the Royal Musical Association*, 126(2):193–249. [Cited on page 212.]

Middendorf, J. and Kalish, A. (1996). The "change-up" in lectures. In *Natl. Teach. Learn. Forum*, volume 5, pages 1–5. Wiley Online Library. [Cited on page 25.]

Nathan, M. J., Kintsch, W., and Young, E. (1992). A theory of algebra-word-problem comprehension and its implications for the design of learning environments. *Cognition and Instruction*, 9(4):329–389. [Cited on page 22.]

O'Neill, B. (2013). *A Computational Model of Suspense for the Augmentation of Intelligent Story Generation*. PhD thesis, Georgia Institute of Technology. [Cited on page 66.]

O'Neill, B. and Riedl, M. (2014). Dramatis: A computational model of suspense. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence*. [Cited on pages 60, 62, and 65.]

O'Neill, B., Riedl, M., Finlayson, M. A., Meister, J. C., and Bruneau, E. G. (2014). Applying qualitative research methods to narrative knowledge engineering. In *5th Workshop on Computational Models of Narrative*, page 139. [Cited on page 65.]

Ortony, A. and Clore, G. (1989). Emotions, moods, and conscious awareness; comment on johnson-laird and oatley's "the language of emotions: An analysis of a semantic field". *Cognition & Emotion*, 3(2):125–137. [Cited on page 33.]

Pérez, R. P. Ý. and Sharples, M. (2001). Mexica: A computer model of a cognitive account of creative writing. *Journal of Experimental & Theoretical Artificial Intelligence*, 13(2):119–139. [Cited on pages 54 and 67.]

Pérez y Pérez, R. (2007). Employing emotions to drive plot generation in a computer-based storyteller. *Cognitive Systems Research*, 8(2):89–109. [Cited on pages 54 and 67.]

Propp, V. I. (1968). *Morphology of the Folktale*, volume 9. University of Texas Press. [Cited on pages 7 and 18.]

Riedl, M. O. and Young, R. M. (2010). Narrative planning: balancing plot and character. *Journal of Artificial Intelligence Research*, 39(1):217–268. [Cited on page 64.]

Rumelhart, D. E. (1975). Notes on a schema for stories. *Representation and understanding: Studies in cognitive science*, 211:236. [Cited on pages 45 and 47.]

Rutledge, R. B., Skandali, N., Dayan, P., and Dolan, R. J. (2014). A computational and neural model of momentary subjective well-being. *Proceedings of the National Academy of Sciences*, 111(33):12252–12257. [Cited on page 200.]

Schank, R. C. and Abelson, R. P. (1975). *Scripts, plans, and knowledge*. Yale University. [Cited on page 30.]

Schank, R. C. and Abelson, R. P. (1977). Scripts, plans, goals and understanding: an inquiry into human knowledge structures. hillsdale, nj: L. [Cited on page 85.]

Schauer, M. and Elbert, T. (2010). Dissociation following traumatic stress. *Zeitschrift für Psychologie/Journal of Psychology*, 218(2):109–127. [Cited on page 4.]

Schnelle, H. and Doust, R. (1988). A net-linguistic "earley" chart-parser. *Connectionist Approaches to Languages, North-Holland, Amsterdam*. [Cited on pages a and 213.]

Schnelle, H. and Doust, R. (1992). A Net-Linguistic "Earley" Parser. In *Connectionist Approaches to Language Processing*, pages 170–205. [Cited on pages a and 213.]

Sgouros, N. M. (1999). Dynamic generation, management and resolution of interactive plots. *Artificial Intelligence*, 107(1):29–62. [Cited on page 50.]

Shklovsky, V. (1965). *Art as technique*. [Cited on page 18.]

Singer, M. and Ferreira, F. (1983). Inferring consequences in story comprehension. *Journal of Verbal Learning and Verbal Behavior*, 22(4):437–448. [Cited on page 26.]

Snowden, D. J. (2000). The art and science of story or 'are you sitting uncomfortably?'. *Business Information Review*, 17(4):215–226. [Cited on page 2.]

Speer, N. K., Zacks, J. M., and Reynolds, J. R. (2007). Human brain activity time-locked to narrative event boundaries. *Psychological Science*, 18(5):449–455. [Cited on page 24.]

Sternberg, M. (2009). Epilogue. how (not) to advance toward the narrative mind. *Applications of Cognitive Linguistics 10*, page 455. [Cited on page 39.]

Sternberg, M. (2010). Narrativity: from objectivist to functional paradigm. *Poetics Today*, 31(3):507–659. [Cited on page 37.]

Sternberg, S., Monsell, S., Knoll, R., and Wright, C. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. *Information processing in motor control and learning*, pages 117–152. [Cited on page 33.]

Stevens, S. S. (1975). *Psychophysics*. Transaction Publishers. [Cited on page 166.]

Sugiyama, M. S. (2001a). Food, foragers, and folklore: The role of narrative in human subsistence. *Evolution and Human Behavior*, 22(4):221–240. [Cited on page 3.]

Sugiyama, M. S. (2001b). Narrative theory and function: Why evolution matters. *Philosophy and Literature*, 25(2):233–250. [Cited on page 3.]

Szilas, N. (2003). Idtension: a narrative engine for interactive drama. In *Proceedings of the Technologies for Interactive Digital Storytelling and Entertainment (TIDSE) Conference*, volume 3, pages 187–203. [Cited on page 56.]

Tan, E. and Diteweg, G. (1996). Suspense, predictive inference, and emotion in film viewing. *Suspense: conceptualizations, theoretical analyses, and empirical explorations*, pages 149–188. [Cited on page 33.]

Tan, E. and Fasting, B. (1996). *Emotion and the structure of narrative film: Film as an emotion machine*. L. Erlbaum Associates. [Cited on page 32.]

Theune, M., Faas, S., Heylen, D., and Nijholt, A. (2003). The virtual storyteller: Story creation by intelligent agents. [Cited on pages 42 and 43.]

Tooby, J. and Cosmides, L. (2001). Does beauty build adapted minds? toward an evolutionary theory of aesthetics, fiction, and the arts. *SubStance*, 30(1):6–27. [Cited on pages 3 and 5.]

Trabasso, T. and Suh, S. (1993). Understanding text: Achieving explanatory coherence through on-line inferences and mental operations in working memory. *Discourse Processes*, 16(1-2):3–34. [Cited on page 26.]

Trabasso, T. and Van Den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of memory and language*, 24(5):612–630. [Cited on pages 47 and 63.]

Trabasso, T., Van den Broek, P., and Suh, S. Y. (1989). Logical necessity and transitivity of causal relations in stories. *Discourse Processes*, 12(1):1–25. [Cited on page 30.]

Truby, J. (2007). *The anatomy of story: 22 steps to becoming a master storyteller*. Faber & Faber. [Cited on page 14.]

Turner, S. (1993). Minstrel: a computer model of creativity and storytelling. [Cited on pages 52 and 67.]

Van Dijk, T. A., Kintsch, W., and Van Dijk, T. A. (1983). *Strategies of discourse comprehension*. Academic Press New York. [Cited on page 22.]

Veldhuis, D. and Kurvers, J. (2012). Offline segmentation and online language processing units: the influence of literacy. *Written language & literacy*, 15(2):165–184. [Cited on page 27.]

Virtanen, T. (2004). *Approaches to cognition through text and discourse*, volume 147. Walter de Gruyter. [Cited on page 93.]

Vorderer, P., Knobloch, S., and Schramm, H. (2001). Does entertainment suffer from interactivity? the impact of watching an interactive tv movie on viewers experience of entertainment. *Media Psychology*, 3(4):343–363. [Cited on page 33.]

Walton, K. L. (1978). Fearing fictions. *The Journal of Philosophy*, pages 5–27. [Cited on page 32.]

Ware, S. G. and Young, R. M. (2014). Glaive: A state-space narrative planner supporting intentionality and conflict. In *Proceedings of the 10th International Conference on Artificial Intelligence and Interactive Digital Entertainment*. [Cited on pages 53, 88, and 105.]

White, W. (1939). *The Psychology of Making Life Interesting*. The Macmillan company. [Cited on pages 20 and 35.]

Wilkens, R. and Schnelle, H. (1990). A connectionist parser for connectionist phrase structure grammars. *Konnektionism in artificial Intelligence and Kognitionsforschung*. [Cited on page a.]

Wilkens, R. and Schnelle, H. (1994). Ansätze zum implementationellen konnektionismus in der kognitiven linguistik. In *Kognitive Linguistik*, pages 9–42. Springer. [Cited on page a.]

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., and Reynolds, J. R. (2007). Event perception: a mind-brain perspective. *Psychological bulletin*, 133(2):273. [Cited on page 23.]

Zhang, L., Barnden, J. A., Hendley, R. J., Lee, M. G., Wallington, A., and Wen, Z. (2008). Affect detection and metaphor in e-drama. *International*

Journal of Continuing Engineering Education and Life Long Learning, 18(2):234–252. [Cited on page 57.]

Zwaan, R. A., Langston, M. C., and Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, pages 292–297. [Cited on pages 24, 30, 81, 150, and 155.]

Zwaan, R. A., Madden, C. J., Yaxley, R. H., and Aveyard, M. E. (2004). Moving words: Dynamic representations in language comprehension. *Cognitive Science*, 28(4):611–619. [Cited on page 32.]

Zwaan, R. A. and Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological bulletin*, 123(2):162. [Cited on page 24.]